THE UNIVERSITY *of* EDINBURGH

The Interpretation of the Crystal Structures of Transition Metal Compounds and Complexes

Matthew G Reeves

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy to the School of Chemistry,

University of Edinburgh

2020

Declaration

I declare that this Thesis has been written by me and that the work carried out is my own, or I have made a substantial contribution towards it except where specific reference is made to another. This work has not been submitted for another degree or professional qualification.

Matthew G. Reeves

List of Publications

1. Reeves, M., Wood, P. and Parsons, S., Automated oxidation-state assignment for metal sites in coordination complexes in the Cambridge Structural Database. *Acta Crystallographica Section B Structural Science, Crystal Engineering and Materials*, **2019** *75* (6), 1096-1105.

2. Reeves, M., Wood, P. and Parsons, S**.** MrPIXEL: Execution of PIXEL Calculations via the MERCURY Interface. *Journal of Applied Crystallography,* **2020**, 53 (4), 1154-1162.

3. Reeves, M., Tailleur, E., Wood, P., Marchivie, M., Chastanet, G., Guionneau, P., and Parsons, S., Revealing Cooperativity Pathways in Spin Crossover Complexes, *In preparation.*

Abstract

The Cambridge Structural Database (CSD) contains over half a million structures containing transition metal compounds. The wealth of data available from these structures is enormous and the potential knowledge and insight that be gained from these structures is of great benefit to the field of inorganic and organometallic chemistry.

Searching through these structures is facilitated by the CSD software Conquest, which allows users to filter results to desired metals and ligand motifs. However, some aspects are not currently possible through the Conquest GUI, while others, such as oxidation state data, require the value to be explicitly mentioned by the author in the assigned chemical name. The introduction of the CSD Python Application Programming Interface (API) has allowed users to interpret, retrieve, and manipulate data in new ways, and now presents an avenue to introduce new data to CSD entries and user structures alike.

A highly reliable automated workflow for oxidation-state assignment in transition-metal coordination complexes has been developed with CSD Python API scripts. These scripts implement the bond-valence sum (BVS) method as well complementary techniques. The strengths and limitations of these methods are discussed and the application of confidence banding for improved assignment confidence is explored. In total, four complementary techniques have been implemented in this study. The resulting workflow overcomes the limitations of the BVS approach, widening the applicability of an automated procedure to more CSD entries. Assignments are successful for 99% of the cases where a high consensus between different methodologies is observed.

Further exploiting the capabilities of the CSD Python API, and specifically the integration with the Mercury software, a program has been developed for the direct execution of PIXEL-CLP calculations from the Mercury interface. The PIXEL method is a semi-empirical procedure for the calculation of intermolecular interaction and lattice energies based on undistorted ab initio molecular electron densities. Following initial set up of a crystallographic model, the ”MrPIXEL” software package assigns atom types and writes necessary input files, submits the required electron density

calculation either locally or to a remote server, downloads the results and submits the PIXEL calculation itself. Full lattice energy calculations can be performed for structures with up to two molecules in the crystallographic asymmetric unit, for more complex cases molecule-molecule energies are calculated only.

Finally, the MrPIXEL software has been used to determine lattice and interaction energies for structures of spin-crossover (SCO) complexes in the CSD. The results of which have been studied in order to determine their influence and role in the abruptness of spin-state transitions for the $Fe(PM\text{-}L)_2(NCS)_2$ family of SCO complexes. The change in interaction energies between spin-states is found to correlate with the abruptness of transition, with more abrupt transitions being associated with much larger changes in interaction energies between spin-states. The interaction energies, along with the changes in energies, are visually displayed in Mercury using a new method developed for producing energy frameworks, similar to those used in CrystalExplorer.

Lay Summary

Within the field of solid-state chemistry, the physical structure of molecules is recorded from x-ray crystallographic data. For molecules which contain carbon, a comprehensive database of these structures is available in the Cambridge Structural Database (CSD), which contains over a million entries. The wealth of knowledge available from data already submitted in the CSD has proved invaluable to the field of organic chemistry, where it enables users to search, retrieve, and analyse current solid structures, particularly in areas such as the pharmaceuticals industry.

The CSD is also home to nearly half a million structures containing transition metal atoms. In comparison to the purely organic structures, the transition metal structures have received less attention in post deposition analysis. This is in part due to aspects of transition metal chemistry remaining missing or underdiscussed in the CSD data.

For instance, the formal charge of the transition metal atom, known as oxidation state, is often missing from uploaded data, making searching and retrieval of relevant structures difficult. To rectify this, new procedures and programs have been developed to add more data to transition metal containing entries within the CSD. This has been facilitated through the Python based Application Programming Interface (API).

Building on this work, new tools have also been developed to study the interactions between molecules in the solid-state. These tools have been used extensively in discussing the physical and chemical properties of organic materials in the past, such as in rationalizing the bioactivity and solubility of drug compounds, though their use in coordination compounds (containing transition metal atoms) has been more limited. By producing new programs that automate computational methods for the calculation of intermolecular interaction energies, it is hoped that their application towards transition metal containing structures will become commonplace.

Finally, and as a demonstration of the benefits of discussing intermolecular interaction energies in the context of coordination complex structures, the software discussed previously has been used to rationalize the magnetic behaviour of a series of coordination complexes.

Acknowledgements

Thanks must first be expressed to Simon Parsons, who, as personal supervisor over this PhD has provided an excellent tutor and mentor, patiently suffering through my terrible grammar and spelling, as well continuing to offer advice and guidance that has been essential to the successes achieved here.

The ever-changing faces of Office 85, whether that be the fantastic Post-Docs, Gary the technician, the undergraduate Master's students, or of course the many great fellow PhD students. All of which have provided fantastic contributions and company which fuelled and aided in many of the topics and results discussed.

To my parents, who have been essential in supporting me through the last four years. Individually, my father, who took it upon himself to understand what I do to a level often beyond that of my own comprehension, discussing the science for many hours and offering staggeringly more input than he realizes. And my mother, who patiently put up with listening to the previously mentioned hours of conversation, completely none the wiser to any of it, suffering from what I can only assume was mind numbing boredom, yet still gladly proof-read the completed article.

To the often overlooked, and generally underappreciated technical and support staff whose advice and hard work kept everything running smoothly and provided facilities key to the results achieved here. Crucially they provided an outlet and a welcome distraction when writing began to overwhelm and infuriate.

A surprisingly large thanks goes to the many undergraduates who offered ideas, input, and support that often proved as, if not more valuable than that of more senior colleagues and staff.

Lastly, I must thank those, regrettably no longer around, who if anything, were perhaps the most instrumental in me getting to this point.

Table of Contents

1. Introduction: The Application of Semi-Empirical Calculations for the Interpretation of Interactions in Crystal Structures

1.1. Thesis Background

Coordination complexes containing transition metals display many properties and characteristics that are of immense interest in commercial and practical applications. In the solid-state for example, variable behaviours in terms of magnetism, colour, and conductivity may be tuned and altered using environmental changes, either through temperature, pressure, or light, potentially offering new materials for use in switching and sensory applications.

Despite the wide-reaching interest in this field, many aspects of the mentioned characteristics are not fully understood, thus making predictions of material behaviour difficult from structural data alone. In addition, this inhibits the use of crystal engineering, the growing field in which new solid structures are designed and retro-synthesized for the properties expected and desired.

Thus, it is essential that further work is done to understand the various properties of coordination complexes, based on structural data currently available. The Cambridge Structural Database (CSD) already contains nearly half a million transition metal containing compounds, providing a significant volume of data ready to be analysed in the quest to determine trends between structural properties and physical characteristics. Indeed Olga Kennard, founder of the Cambridge Structural Database (CSD), stated that:

"*collective use of data would lead to the discovery of new knowledge which transcends the results of individual experiments*"[1]

However, while the volume of information within the CSD continues to grow at a near exponential rate, many structures are lacking key information that is otherwise essential to the efficient searching and filtering of data. For transition metal complexes, this issue is most noticeable in terms of oxidation state information. In 2016, oxidation state data was missing from over half the entries deposited (Figure 1.1).

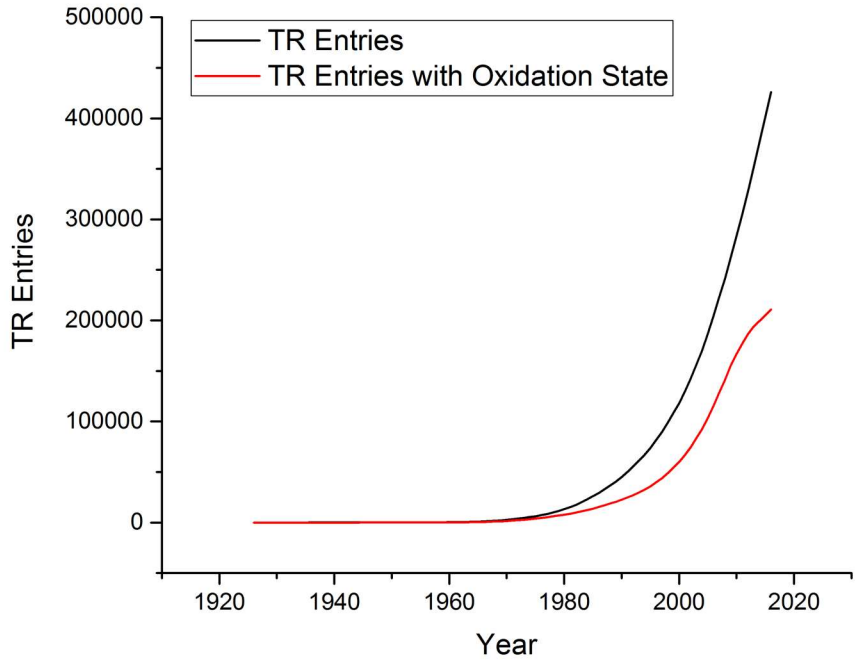


Figure 1.1: Number of transition metal containing entries in the CSD by year in black with the number of those entries containing oxidation state data in red. Data from 2016.

This problem is compounded by the restrictions of the current CSD searching system. Structure retrieval is currently performed using the CSD software Conquest, which allows a user to search for structures using a variety of approaches, which may be combined, as necessary. This includes drawing molecular fragments, inputting chemical names or formulas, or restricting searches to specific crystallographic settings. While the array of search parameters provides an excellent set of tools for retrieving relevant and desired organic structures, they are lacking certain features which are of major benefit when looking for transition metal structures. Oxidation states for example, cannot be explicitly defined in the search parameters, and instead must be defined only by their presence (in roman numeric form) in the chemical name. E.g. "must contain in the chemical name: iron(iii)" when searching for iron 3+ structures.

While this issue urgently needs addressing to improve the efficiency and ease of retrieving relevant structures, the structural data is already present for further analysis to take place. For the crystalline structures of molecular solids, this focuses on the intermolecular interactions between the ordered arrangement of molecules.

1.2. Intermolecular Interaction Energies

1.2.1. Intermolecular Interactions

In crystalline solids, the ordered arrangement of intermolecular interactions is a consequence of the geometry and electronic structure of the individual ions, atoms or molecules. The strength and vibrational characteristics of these interactions defines the physical characteristics of the solid, as well as structural changes that occur when the material is subjected to changes in temperature and pressure. Intermolecular interactions are invoked when interpreting and analysing crystal structures in the context of their stability and physical properties. For example, in pharmaceuticals, the intermolecular forces are considered in relation to solubility and bioavailability, as well as in the formulation of co-crystals with optimum properties for tabletting. In coordination chemistry, these interactions influence electronic properties such as magnetism. By considering intermolecular interactions in their role in structure formation, the rapidly growing field of crystal engineering aims to synthesize new structures with desired properties.

Provided that no covalent effects are involved, intermolecular interactions can be considered in terms of four fundamental energy contributions that describe the various forms of stabilizing and destabilizing forces that occur. These energy terms arise from electrostatic (or Coulombic), polarisation (or induction), dispersion (or van der Waals or London) and Pauli (or short-range) repulsion effects.

1.1.1. Electrostatic interactions

Electrostatic interactions are caused by the attractive forces of opposing charges and repulsive forces of like charges. The influence of these forces, the electrostatic energy (E_{Elec}), is expressed by Coulomb's Law:

$$E_{Elec} = \frac{q_1 q_2}{4\pi\varepsilon_o r_{12}} \qquad 1.1$$

Where q_1 and q_2 are the point charges separated by distance r_{12} and ε_0 is the vacuum permeability. Where the signs of q_1 and q_2 are the same, E_{Elec} is positive, and repulsive, while opposite signs result in a negative E_{Elec} and a stabilizing contact.

In molecular systems electrostatic interactions are caused by the variation of the local charge density distribution in a molecule, which forms positive (δ+) or negative (δ−) regions depending on the electronegativity of the atoms present and the overall charge if the species is an ion. Interactions that are predominantly electrostatic in nature are often identifiable by the proximity of atoms with significantly different electronegativities, as in the case for hydrogen-bonding illustrated in Figure 1.2 by interactions in the crystal structure of γ-glycine.

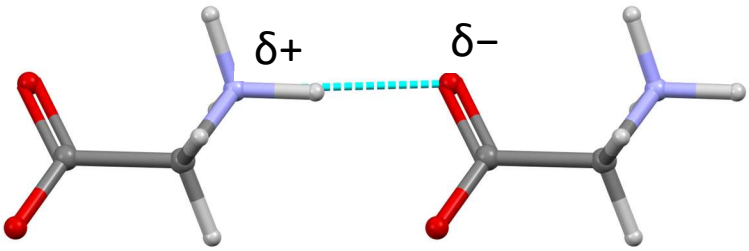


Figure 1.2: Example of an electrostatic dominant interaction in γ-glycine between δ− oxygen atom and δ+ RNH_3^+ proton.[2]

1.1.2. Polarisation Interactions

When a point charge approaches a spherical cloud of electron density around an atom, the electron density will distort to produce a stabilising polarisation of the charge (Figure 1.3 i. and ii.). Polarisation interactions are always stabilising. The magnitude of the polarisation energy (E_{Pol}) is dependent on distance (varying as r^{-6}), the approaching charge and the polarizability (α) of the electron cloud, which measures its deformability. In molecular interactions this relationship is expressed by Equation 1.2 where μ is the dipole moment of an approaching molecule, and r_{12} is the separation between molecules 1 and 2.

$$E_{Pol} \propto -\frac{\alpha\mu^2}{r_{12}^6} \qquad 1.2$$

The value of the polarizability depends on the distance from the electrons from the nucleus. Larger atoms will have larger polarizabilities, allowing for a greater deformation of electron density, whereas smaller atoms, in which the nuclear attraction is higher, will be less susceptible to approaching point charges and have a lower polarizability, illustrated for the noble gases in Figure *1.3* iii. In practice, the

contribution of the polarisation term in molecular crystal structures is often smaller than other terms. In the hydrogen bond formed between the glycine molecules shown in Figure *1.2*, the electrostatic energy is −105.8 kJ mol^{-1}, while the polarisation contribution is −7.6 kJ mol^{-1}, as determined by CrystalExplorer.[3]

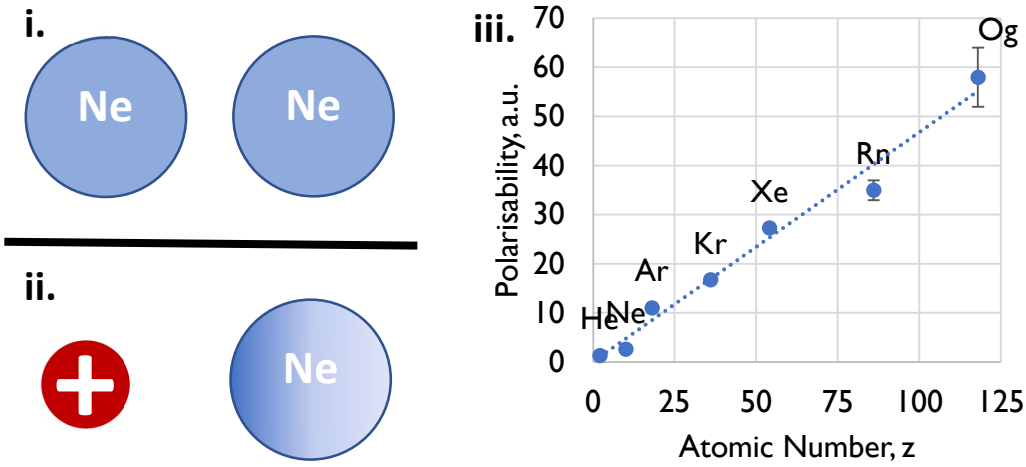


Figure 1.3: i. Neon atom electron clouds as spheres with no introduced external charge. ii. Neon cloud deformation with introduction of positive point charge. Electron Density shift denoted by colour gradient. iii. Increase in polarizability with increasing atomic size for noble gasses.

1.1.3. Dispersion Interactions

While polarisation interactions are produced by permanent dipole moments in neighbouring molecules, instantaneous fluctuations in electron density also occur which produce instantaneous atomic and molecular dipoles. This polarisation of electron density causes further deformation of electron density in neighbouring molecules. As with polarisation interactions, these are proportional to the polarizability of the atoms involved and their separation.

As was the case for polarisation, dispersion interactions are always stabilising, and also vary as r^{-6}, declining rapidly with distance. The dispersion energy, E_{Disp}, is approximated with the London Dispersion Equation:

$$E_{Disp} \approx -\frac{3}{2}\frac{I_1 I_2}{(I_1 + I_2)}\frac{\alpha_1 \alpha_2}{r_{12}^6} \qquad 1.3$$

The ionisation energies of the interacting atoms (I_1 and I_2) were introduced by London to represent the frequency of charge oscillations. The dependence on polarisabilities mean that E_{Disp} increases with increasing atomic size and electron density.

Dispersion interactions are maximised when strongly polarisable groups lie in close proximity to one another in crystal structures, optimising a large contact surface between polarisable regions of electron density. As is the case with aromatic ring (π-π) stacking, which can be identified by inspection of a crystal structure packing diagram. Other dispersion interactions lack a characteristic geometrical signature, and this has sometimes led to their importance remaining unrecognised. By contrast to electrostatic interactions, they are often better understood in terms of whole molecule – whole molecule (as opposed to atom-atom) contacts which take into account the overall topology of the molecular contact surface.

Electrostatics and dispersion are usually found to be the dominating stabilising terms in molecular crystal structures, although the former generally dictate the orientation of molecules within the structure rather than the latter. This can be exemplified using Hirshfeld surfaces of the γ-glycine example, shown in Figure 1.2, where the surfaces are coloured according to electrostatic potential (red is negative, blue is positive). Figure 1.4 shows how the packing of molecules is arranged such as to maximize the contact between opposingly charged areas of each molecule. This is also the case for molecules with dispersion dominant interactions, such as in the crystal structure of benzene, where the packing of molecules is arranged such to optimise electrostatic interaction, with the δ+ hydrogen atoms pointing into the centre of δ− aromatic rings.

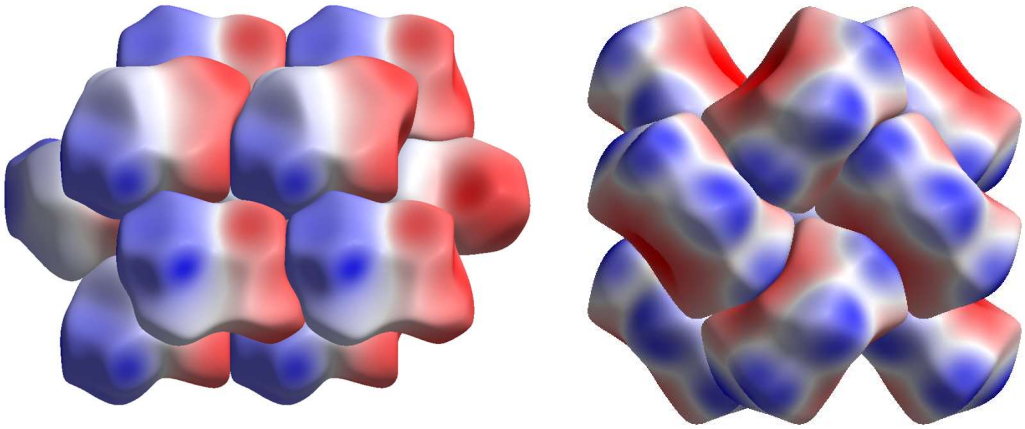


Figure 1.4: Hirshfeld Surfaces with electrostatic potentials mapped for first coordination sphere of γ-glycine (left) and benzene (right). Positive regions are blue, negative regions red. A full description of Hirshfeld Surfaces is available in Section 1.3.3.[2,4]

1.1.4. Repulsion Interactions

The Pauli repulsion term is quantum mechanical in origin, is always destabilising, and diminishes even more rapidly with distance than the polarisation and dispersion energies. The Pauli Exclusion Principle forbids the overlap of electrons with the same spin state occupying the same space. This differs from the electrostatic term, which is charge based repulsion of the corresponding electron clouds. As two atoms or molecules approach each other their electron clouds begin to overlap. As a result, the electron density is deformed away from the contact partially exposing the positively-charged nuclei leading to development of a destabilizing energy, E_{Rep}.

The repulsion energy varies as r^{-12}, but beyond this it is difficult to calculate accurately, and it is often approximated with an empirical relationship, such as that shown in Equation 1.4, where it is made proportional to a constant D.

$$E_{Rep} = \frac{D_{12}}{r_{12}^{12}} \qquad 1.4$$

Pauli and electrostatic repulsion are the only destabilising terms in the model of intermolecular interactions described here. Within a first molecular coordination sphere both can combine to make some interactions destabilising overall, with the influence of Pauli repulsion becoming very important at high pressure. This can be

seen for glycine (Figure 1.5) where repulsion terms increase significantly with increasing pressure. Their contribution to the total interaction energies can be shown with the shortest contact. At ambient pressure, the shortest contact in α-glycine has a total interaction energy of -166.9 kJ mol^{-1}, and a repulsion term of 40.8 kJ mol^{-1}. For the structure at 6 GPa, the total interaction energy is −183.2 kJ mol^{-1}, with a repulsion term of 78.0 kJ mol^{-1}. However, the sharp decline of the repulsion term with distance means that beyond the first coordination sphere there are only electrostatic interactions, which vary as r^{-1}, remaining to balance other stabilising terms.

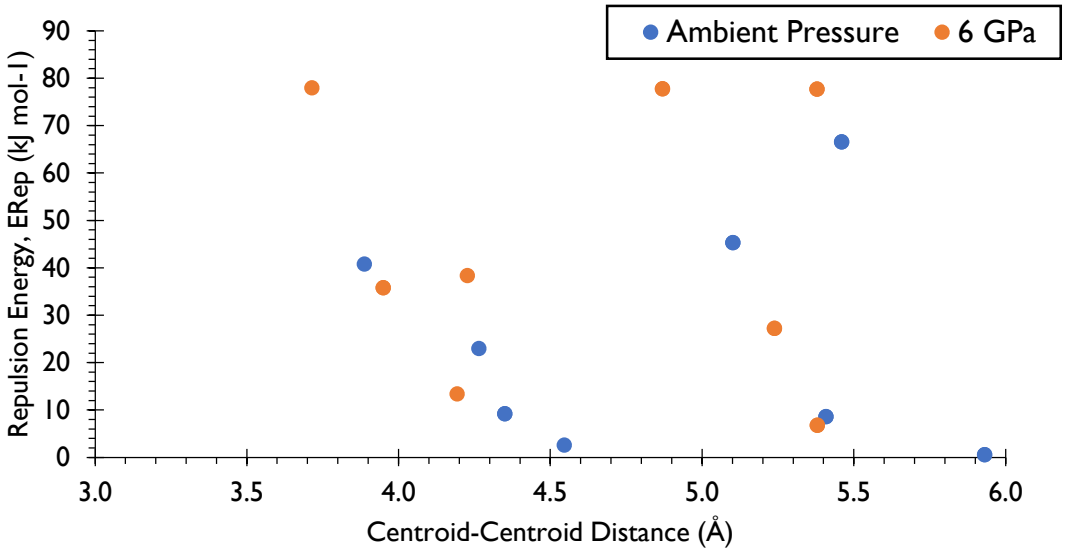


Figure 1.5: Repulsion energy term for short interactions in α-glycine with increasing pressure. Values determined using the PIXEL-CLP method.[5,6]

1.1.5. Example: Intermolecular Interactions in γ-Glycine

The use of the four terms described above in practical crystal structure analysis can be illustrated with γ-glycine (Figure 1.6). Interactions for which the Pauli repulsion energy is greater than zero identify the first molecular coordination sphere, which contains seven pairs of crystallographically unique interactions. The total coordination number is fourteen, making the structure topologically related to body-centred cubic packing of hard spheres. The interaction energies can be calculated using many of the methods discussed later in this Chapter, but here they were obtained using CrystalExplorer.

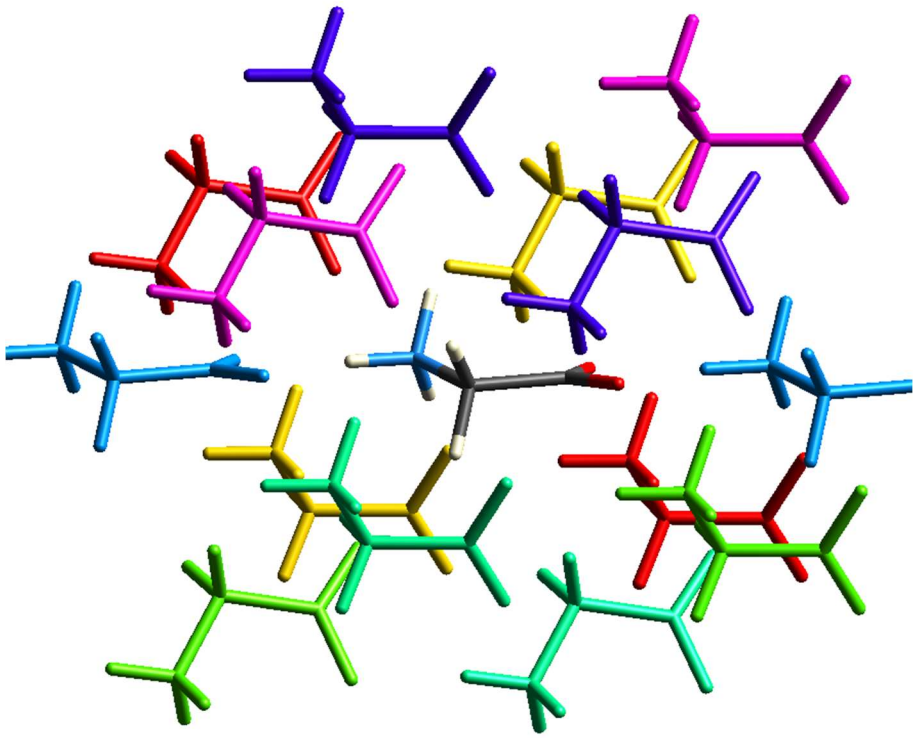


Figure 1.6: First coordination sphere of γ-glycine with unique interactions to a central molecule coloured in CrystalExplorer. A breakdown of the calculated interaction energies is available in Table 1.1.

Table 1.1: Breakdown of interaction energies as calculated by CrystalExplorer for first coordination sphere of γ-glycine. Colours match molecules in Figure 1.6. Note that due to crystal symmetry, each contact has an equivalent (as grouped by colours).

Colour Indicator	***Sym. Op.***	***R***	E_{Elec}	E_{Pol}	E_{Dis}	E_{Rep}	E_{Tot}
		(Å)			kJ mol^{-1}		
	x, y, z+1	5.47	−105.8	−27.6	−9.6	56.2	−105.9
	x,y,z−1	5.47	−105.8	−27.6	−9.6	56.2	−105.9
	−y,−1+x−y,2/3+z	5.37	−35.8	−12.5	−7.7	8.8	−48.4
	1−x+y,−x,−2/3+z	5.37	−35.8	−12.5	−7.7	8.8	−48.4
	−x+y,−x,−2/3+z	5.64	−23.1	−12.0	−6.7	8.6	−33 .9
	−y,x−y,2/3+z	5.64	−23.1	−12.0	−6.7	8.6	−33 .9
	1−x+y,1−x,1/3+z	4.27	−23.7	−33.9	−14.6	51.9	−30.8
	1−y,x−y,−1/3+z	4.27	−23.7	−33.9	−14.6	51.9	−30.8
	1−x+y,1−x,−2/3+z	5.31	9.8	−5.2	−3.9	1.4	4.0
	1−y,x−y,2/3+z	5.31	9.8	−5.2	−3.9	1.4	4.0
	−y,x−y,−1/3+z	4.68	29.9	−30.3	−13.8	35.3	19.1
	−x+y,−x,1/3+z	4.68	29.9	−30.3	−13.8	35.3	19.1
	−y,−1+x−y,−1/3+z	4.34	47.1	−8.8	−11.1	6.7	37.8
	1−x+y,−x,1/3+z	4.34	47.1	−8.8	−11.1	6.7	37.8

Of the seven unique contacts in the first coordination sphere, four are stabilizing and three are destabilising, showing the assumption that short intermolecular interactions always stabilise a crystal structure can be incorrect. The source of the destabilisation is usually electrostatic repulsion, but in one case Pauli repulsion is the dominating term.

In most cases the electrostatic term is the largest term, as would be expected in the crystal structure of a zwitterionic species such as glycine. Indeed, hydrogen bonds are present in all four of the stabilizing interactions. In particular, the strongest interaction (translations along the *c* axis, −105.9 kJ mol^{-1}) are those that contain a short charge-assisted NH···O hydrogen bond, which contributes to the dominant electrostatic term (−105.8 kJ mol^{-1}).

The presence of hydrogen bonding cannot be considered indicative of a stabilizing molecule-molecule contact however, as the second most destabilizing interactions (19.1 kJ mol^{-1}) also contain a short NH_3···O distance with ideal hydrogen bonding parameters (NH···O = 1.95 Å, <NH···O =178.6°, <CO…H = 135.4°). The contact is destabilizing overall due to the proximity of neighbouring NH_3^+ groups.

The most repulsive interaction occurs between molecules in the same orientation and is dominated by a large electrostatic term (47.1 kJ mol^{-1}). This is in part due to short contact distances between neighbouring COO^- groups, and neighbouring NH_3^+ groups.

1.3. Intermolecular Interactions with Distance

Equations 1.1-1.4 show that inter-atomic and -molecular distances affect the magnitudes of each energy term. At an infinite distance the interaction energy between two molecules is 0. As the distance decreases, the energy tends towards an energy minimum, before the repulsion energy term dominates at short-range. This relationship is modelled with the Lennard-Jones potential (V_{LJ}, Equation 1.5).[7]

$$V_{LJ} = \varepsilon\left[\left(\frac{r_m}{r}\right)^{12} - 2\left(\frac{r_m}{r}\right)^{6}\right] \qquad 1.5$$

where r_m is the optimal interaction distance (energy minimum), r is the separation, and ε is the depth of the potential energy well. Equation 1.5 describes an anharmonic curve where the energy turns significantly positive and repulsive when moving closer than the minimum, while gradually tailing off to 0 with increasing separation away from the minimum (Figure 1.7).

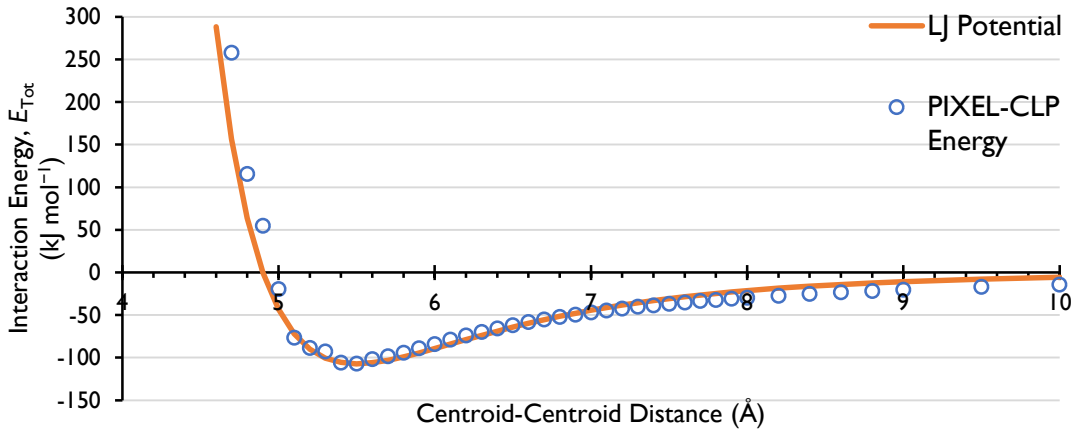


Figure 1.7: Typical anharmonic curve as seen for a γ-glycine contact in Figure 1.2.[2] The curve follows the expression of Equation1.5. The depth of the energy well, ε = 107.2 kJ mol^{-1} and optimal separation r_m = 5.5 Å. The blue points are intermolecular interaction energies as determined by PIXEL-CLP.[8]

Anharmonicity applies to all instances where an attractive force is present in both for atomic and intermolecular interactions (i.e. all bound interactions). In addition to the optimal distance and energy of the interaction, which are given by the position of the minimum, the curvature of the function (second derivative of the curve) defines the rigidity, or resistance to change in distance. This information yields the force constant and hence the vibrational frequency, a steep curve resulting in a smaller vibrational amplitude, and a higher vibrational frequency.

1.3.1. Analysis of Intermolecular Interactions using Distance

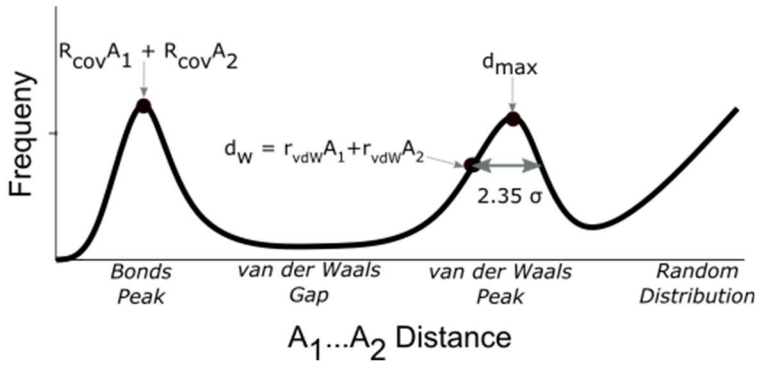
The anharmonic relationship between distance and interaction strength discussed above is key to the use of van der Waals (vdW) radii in interpreting intermolecular interactions. Van der Waals radii are used to define the typical separation between atoms, supposing each atom as a sphere that may overlap another atomic sphere, and thus interact strongly. The curve from Figure 1.7 has an energy minimum at a certain separation between atoms (r_m). For the noble gasses, i.e. non-metallic monoatomic structures, the van der Waals radius, r_{vdW}, is equal to ½ r_m, while for metals ½ r_m represents the metallic radius, and for a homonuclear single covalent bond it represents the covalent radius.

For an intermolecular contact involving two atoms, an indication of the strength of the interaction can be obtained by comparing the contact distance with the sum of van der Waals radii of the two atoms. Where the contact is shorter than the sum of the van der Waals radii the energetic interaction between atoms is considered to be stabilising, whereas beyond this distance the interaction is taken to be weak. This viewpoint is strongly argued by Desiraju, who suggests that in the majority of situations where short-contacts are observed, the interaction is both stabilizing and instrumental in the structural arrangement.[9] Thus contacts with distances within the sum of the van der Waals radii are considered to 'direct' the packing arrangement of the crystal structure.

Values for the van der Waals radii of atoms have been determined empirically from crystal structures, most notably by Bondi, whose tabulated values dating from 1964 still form the basis of many structural analyses today.[10] More recent studies have extended values of van der Waals radii to the whole periodic table, for example Alvarez in 2013 derived 99 values fitted to structural data across the whole Cambridge Database.[11]

Alvarez determined the distribution of interatomic distances for different element-oxygen pairs, finding a general trend (Figure 1.8i) consisting of two peaks, illustrated for Sr···O interactions in Figure 1.8ii.

i.



ii.

Figure 1.8: i. General trend of interatomic distances for atom pairs A_1 and A_2 in crystal structure data, ii. Distribution of Sr···O interatomic distances in CSD structures. Figures reproduced from study by Alvarez. *et al.*[11]

The first peak corresponds to the sum of covalent radii ($R_{cov}A_1 + R_{cov}A_2$) at around 2.8 Å, followed by a distance range devoid of contacts, termed the van der Waals gap, and then a second peak for van der Waals contacts. Beyond the second peak, a random distribution is found with increasing distance for non-interacting contacts. The second peak, d_{max} is comprised of a mixture of van der Waals interactions and interactions from the tail of the random distribution. The sum of van der Waals radii, d_w ($r_{vdW}A_1 + r_{vdW}A_2$), is defined at half the peak height of d_{max}, assuming a Gaussian distribution, where σ is the width of the peak at half-height. The van der Waals radius for oxygen was determined from O···O contacts as 1.50 Å, slightly shorter than

Bondi's value of 1.52 Å. This acts as a reference value which can be used to obtain the van der Waals radii for all other elements when applied to other element···O contacts. For the Sr···O example discussed above, d_w = 4.34 Å, therefore the van der Waals radius for Sr: 4.34 Å – 1.50 Å = 2.84 Å (Equation 1.6).

$$rA_1 = d_{w(A_1 \ldots A_2)} - rA_2 \qquad 1.6$$

The use of van der Waals in interpreting crystal structure packing provides a quick assessment of intermolecular contacts in the context of other, similar, interactions. A visual indication of van der Waals contacts can be produced in the CSD program Mercury (
Figure 1.9).

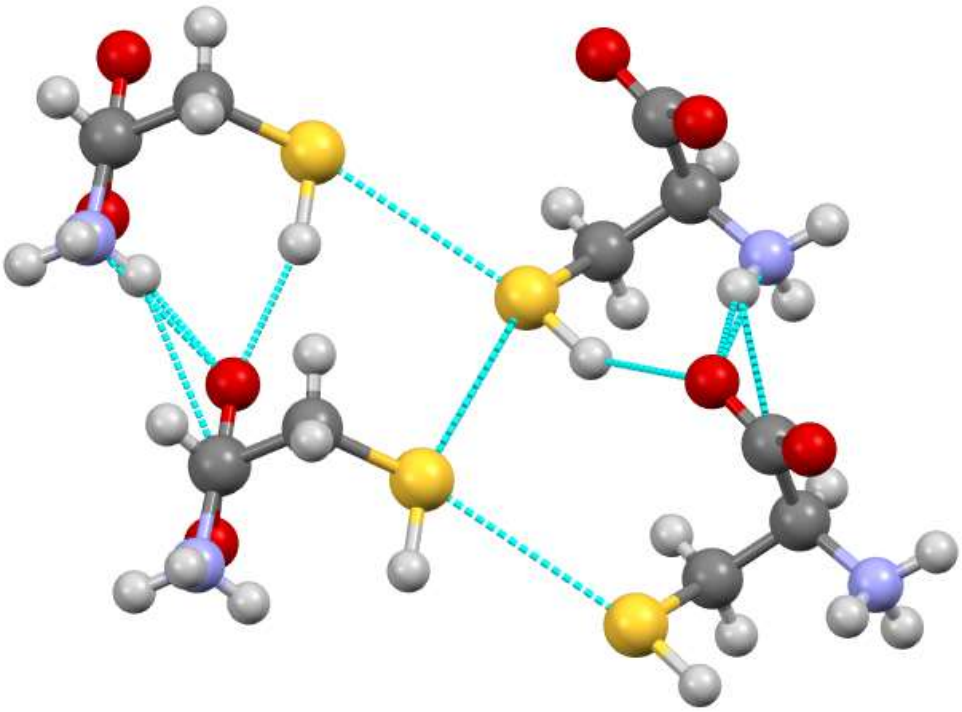


Figure 1.9: Short S···S contacts in L-Cysteine polymorph-II with atom-atom contacts shorter than sum of van der Waals radii shown as blue tie-lines within CSD Mercury program.[12]

For example, if a structure contains a S···S contact of 3.58 Å, then the information that $2xr_{vdW}(S)$ = 3.6 Å shows quickly that this is likely to be structurally important without needing to carry out geometric analysis of similar interactions.

Although it is fast and convenient, there are some limitations and weaknesses in using van der Waals radii as the sole method to interpret crystal structures. Firstly, it works best for interactions based on specific atom-atom contacts, such as O···H hydrogen bonds and aromatic ring stacking, but neglects contacts which are better understood as whole-molecule – whole-molecule interactions. Indeed, some structures occur without any contacts shorter than the sum of van der Waals radii at all. Figure 1.10 shows the relative frequencies for the number of intermolecular contacts formed by small organic molecules within the CSD. There is a distribution with a maximum at 5 contacts, but a sizable number of structures contain no contacts (5,771 structures, 3%).

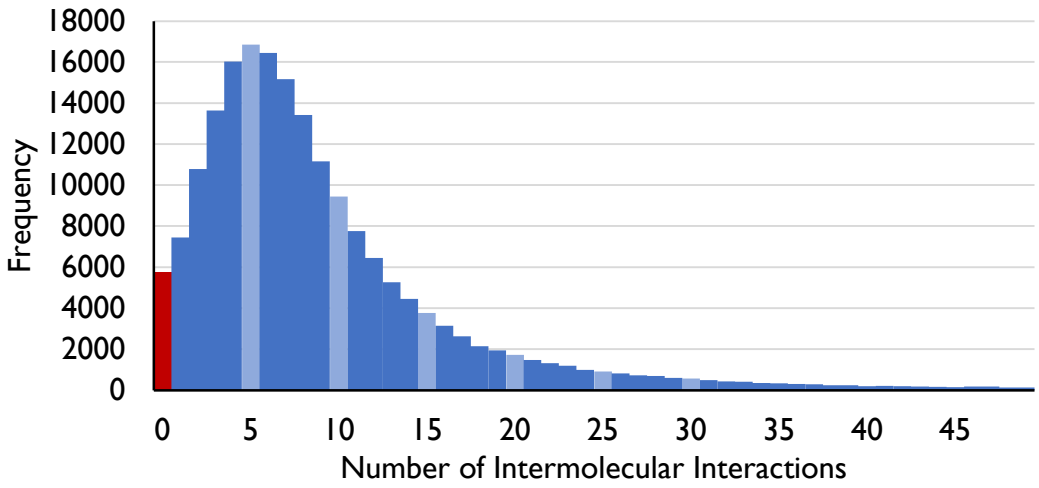


Figure 1.10: Distribution of number of intermolecular contacts for small (< 50 atoms) organic molecules within the CSD. Red bar highlights structures without any contacts shorter than van der Waals radii.

Secondly, analyses based on van der Waals radii give no quantification of the relative strengths of interactions in a structure. This can be seen from the γ-glycine structure discussed previously, where five short contacts are identifiable as hydrogen-bonds that, based on distances alone, would suggest a series of similar stabilising interactions. In contrast, computational techniques reveal very different energies for these contacts, [1] including, as discussed above, one destabilising interaction, ranging from −105.8 to +19.1 kJ mol^{-1}.

Finally, van der Waals are anisotropic, Nyburg *et al*. showed that the distance of the van der Waals radius around an atom is not always constant and therefore instead of being considered spherical should be thought of as elliptical in shape.[13] This is

particularly true for halogen atoms, where, when bonded to carbon, have a significantly shorter van der Waals radii along the C-X vector (known as polar-flattening).

In the light of these problems, there is growing interest in evaluating intermolecular energies in crystal structures. These methods range from the rigorous but slow application of quantum mechanics, to the semiempirical methods such as molecular mechanics, which are very fast but may be limited in applicability to a range of related materials.

1.3.2. The Bond-Valence-Sum Method

The analysis of crystal structures based on interatomic distances can be taken further by incorporating concepts from Pauling's rules for the rationalization of ionic structures.[14] Originally derived from the ionic model, the Bond Valence-Sum method for crystal structure analysis determines relative bonding strengths from interatomic distances and empirical parameter values for specific atomic pairs. A summation of bond valences, or Bond Valence Sum (BVS) around a given atom gives the atomic oxidation state (Equation 1.7).

$$V_{Atom} = \sum v_i \qquad 1.7$$

where V_{Atom} is the valence of the atom in question, and v_i corresponds to the valence of each bond formed around the central atom. Brown *et al.* derive a two-parameter approximation for the experimental bond strength (Equation 1.8).[15]

$$S_{ij} = exp\left(\frac{R_o - R_{ij}}{B}\right) \qquad 1.8$$

Where R_{ij} is the bond length, and R_0 and B are the parameters specific to the cations and anions at a given charge. The values of R_0 and B are tabulated from experimental structure data, with B typically a constant of 0.37 Å, although larger values are determined for more polarisable atoms.[15] Bond Valence Sums, as defined by Equation 1.8, can be applied out to any distance from the central atom, although in practice values drop to zero between 3 and 4 Å, dependant on the atoms present. Therefore, BVS calculations typically focus on the first intramolecular coordination sphere around an atom.

The most common use of the BVS method is in validating and interpreting inorganic and metal-organic structures with respect to the intramolecular bonds of metal centres. This is often exploited in the determination of metal oxidation states. However, the BVS method may be used in interpreting specific contacts within crystal structures. Like the van der Waals model, the BVS model proposes that short contacts are stronger than long ones. By determining the specific valence of the contact using the interatomic distance (R_{ij}) and parameters (R_0 and B) in Equation 1.8, valences are produced which are comparable between different bonding pairs. This approach is employed when looking at perovskite ABX_3 structures, where relative bond strengths calculated from BVS can be used to compare AX and BX contacts within the structure. For example, in the perovskite $Ba(Mn_{½}Ti_{½})O_3$, BVS was used to determine the preferential sites for the Mn and Ti ions, supposing higher bond valences equate to preferable positions.[16] Table 1.2 shows the preferential site for both metals is M(1) due to the highest valences calculated. Ti-O bonds produce higher valences for all sites and thus it is suggested that the M(1) site is occupied preferentially by Mn atoms, while the M(2) and M(3) sites are occupied by Ti atoms.

Table 1.2: Calculated bond valences for Mn-O and Ti-O at metal sites M1-M3 in $Ba(Mn_{½}Ti_{½})O_3$ structure.[16]

	Metal Site		
	M(1)	M(2)	M(3)
S_{Mn-O}	3.98	3.49	3.24
S_{Ti-O}	4.70	4.13	3.83

However, there are some issues with BVS. The correlation between bond valence and bond length, as defined by 1.8, fits an exponential curve, becoming much steeper with decreasing bond distance. Therefore with increasing cation valence, and thus shortening bond lengths, the bond valences are much more sensitive to small changes in length with increasing atomic charges and decreasing bond lengths. Figure 1.11 shows the distribution of Mn-O bond lengths in the CSD for Mn(2+), Mn(3+) and Mn(4+). As expected, the bond lengths decrease with increasing charge. The lower plot shows the effect of bond length on calculated bond valence using the parameters for each oxidation state. The curves are very similar between oxidation states,

showing a steepening curve with decreasing bond distance. Therefore for the higher oxidation states where bonds are typically shorter, the calculated bond valence is considerably more sensitive to the quality of crystal structure and the accuracy of the atomic positions in question.[17,18]

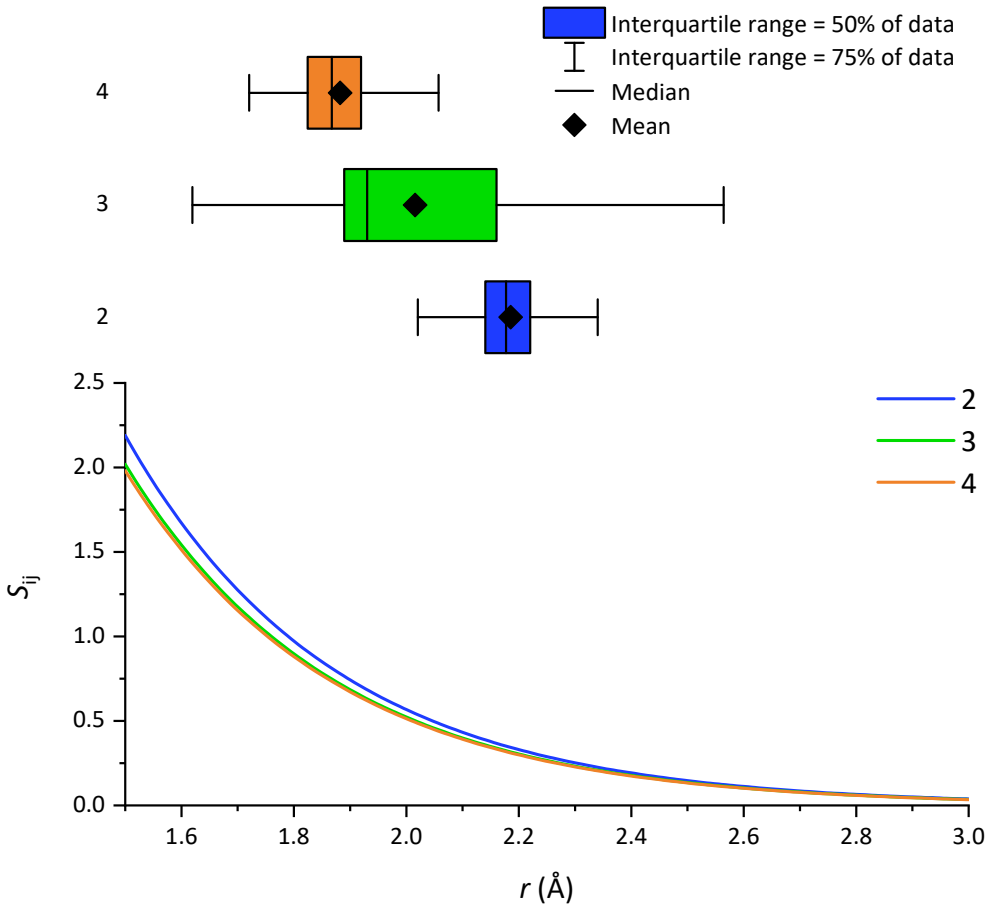


Figure 1.11: (top), Box plot for distribution of Mn-O bond lengths in the CSD by Mn valence, and (bottom), influence of Mn-O bond length on bond valence, S_{ij} as determined using the parameters R_0 and B compiled by Brown.[19] Note slightly different curves for different valences due to the subtle change in the parameter R_0 between Mn^{2+} (blue), Mn^{3+} (green) and Mn^{4+}(orange). Boxes represent lower and upper quartiles (Q1-Q3 (50% of data), while whiskers represent range encompassing 75% of data.

1.3.3. CrystalExplorer for Structural Analysis

The CrystalExplorer software developed by Spackman *et al.* provides a suite of novel tools for interpreting interactions within crystal structures.[3] The interactions are visualized by mapping various parameters to Hirshfeld Surfaces so that prominent contacts can be readily identified from the structure displayed. The Hirshfeld Partitioning Scheme generates a surface around a molecule at a boundary point between monomers. The surface is plotted where the conditions are satisfied:

$$\left|\frac{r_A}{R_A}\right| = \left|\frac{r_B}{R_B}\right| \qquad 1.9$$

where r_A/r_B and R_A/R_B are the vector distances and van der Waals radii respectively between atoms A and B in two molecules. The result is a 3D globular shape that aids in both describing the packing of molecules in the solid structure, and the specific contacts observed (Figure 1.12). The specific contacts between molecules can be visualized and described using multiple surface property options.

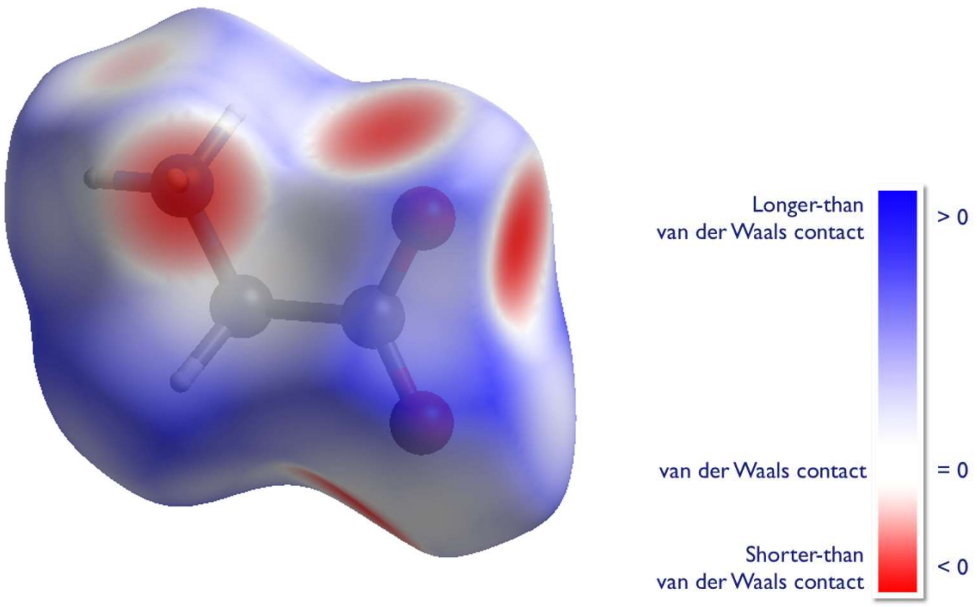


Figure 1.12: Hirshfeld Surface generated by CrystalExplorer for γ-glycine using default settings. Coloured shading represents d_{norm} value. which highlights the contacts between molecules.

This is often determined using the normalized contact distance (d_{norm}) value which displays where molecules become closer than the van der Waals radii separation of molecular atoms (Equation 1.10).

$$d_{norm} = \frac{d_i - r_i^{vdW}}{r_i^{vdW}} + \frac{d_e - r_e^{vdW}}{r_e^{vdW}} \qquad 1.10$$

Where d_i is the distance between the nearest atomic nucleus, *i*, and the Hirshfeld surface and d_e is the distance between the surface and the nearest atomic nucleus, e, in an adjacent molecule, and r_i^{vdW} and r_e^{vdW} are the van der Waals radii of atoms *i* and e respectively. This geometrical analysis can be enhanced by utilising ab initio or semi-empirical calculations executed with Gaussian or Tonto.[20,21] This allows for the mapping of various electronic properties such as the electrostatic potential (Figure 1.13).

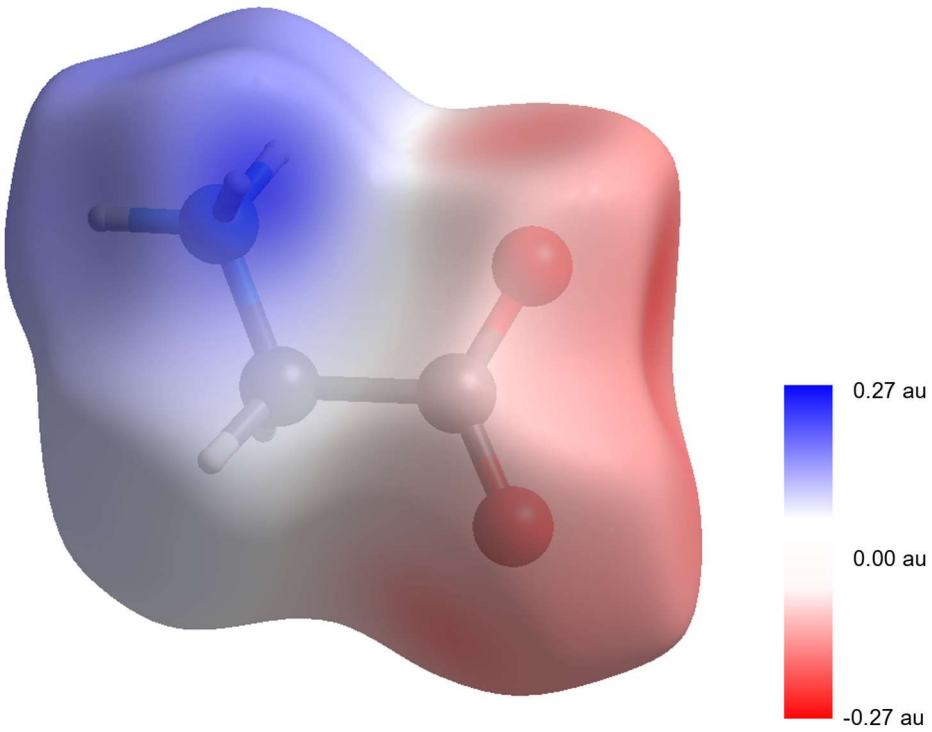


Figure 1.13: Electrostatic Potential mapped onto a Hirshfeld Surface of γ-glycine in CrystalExplorer. Energy calculations carried out with Gaussian using HF-STO-3G level of theory.

With this data it is possible to interpret the nature and packing of molecules within the crystal structure using computational electronic values over simple geometric data with visual results. Therefore one can appreciate the influence of factors such as molecular charge distribution in the formation of crystal structures and in specific contact orientations (Figure 1.14).

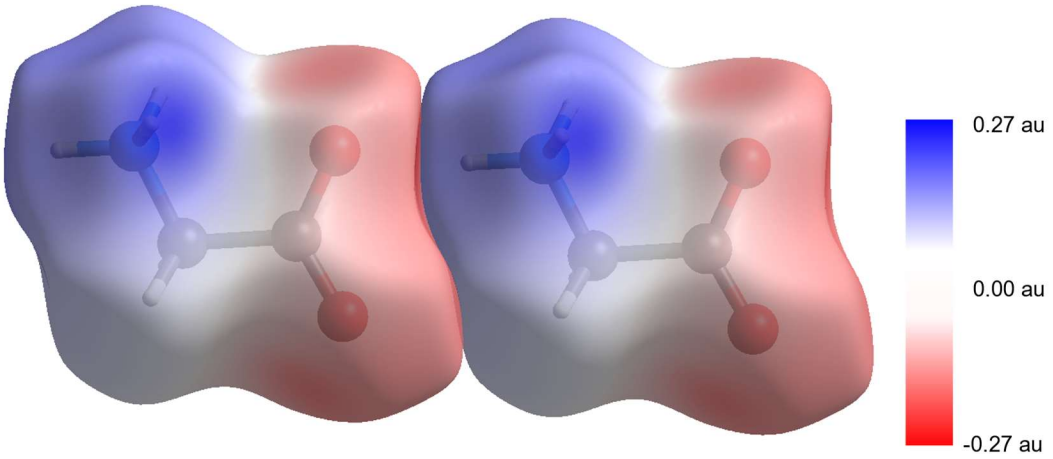


Figure 1.14: Packing of γ-glycine as shown in Figure 1.2 with Hirshfeld Surfaces. Shading relates to the electrostatic potential and demonstrates the optimal contact orientation between opposing charges.

1.4. Analysis of Intermolecular Interactions using Computational Methods

1.4.1. Traditional Quantum-Mechanical Approaches to Intermolecular Energies

In principle, the interaction energy for a molecular dimer, such as that shown in the example for glycine in Figure *1.2*, can be evaluated straightforwardly using a quantum mechanical method such as Møller–Plesset (MP) perturbation theory or Density Functional Theory (DFT). One or two calculations are carried-out to evaluate the energy or energies of the monomer of a homo-molecular contact ($E_A + E_A$) or monomers for a hetero-molecular contact ($E_A + E_B$). A further calculation is carried out for energy of the dimer (E_{dimer}). The interaction energy, E_{int} is then given by Equation 1.11.

$$E_{int} = E_{dimer} - E_A - E_B \qquad 1.11$$

There are a number of caveats that need to be applied to this simple procedure. The first is that density functional methods do not account for dispersion interactions.[22,23] Since these are very important in intermolecular interactions, this is a serious

shortcoming. In 'DFT-D' methods, an empirical dispersion term similar to that used in molecular mechanics is simply added to the DFT energy. This simple solution has been remarkably successful as shown by Grimme *et al.*[24-27] These studies showed that results comparable to the S22 dataset (see below) could be produced, with the method also being applicable to larger macromolecular structures.

A second problem arises when the dimer is taken from a crystal structure. Small experimental measurement errors can have a larger effect on the calculated energies than the dimerization energy itself. This is a particular problem for compounds containing hydrogen, the positions of which are systematically in error when determined from X-ray data. The problem can be reduced by optimising the geometry of the crystal structure by periodic DFT methods and carrying out the calculation on this structure rather than the experimental one.

Finally, the calculations themselves also produce erroneously more positive monomer energies, which erroneously produce more stable dimerization energies. This is due to the larger number of basis functions employed for the dimer, which enables more flexibility and therefore a lower energy. The monomer-dimer energy difference is thus over-estimated, an effect known as known as the Basis Set Superposition Error (BSSE).[28-30] To overcome this, empty "ghost-orbitals" may be added to the monomer calculation basis sets. In this way, comparative values from the monomer and dimer calculation are produced and the difference between monomer/dimer energies is closer to the interaction energies expected by higher level calculations.

Table 1.3: S22 dataset of calculated dimer energies. SAPT and CLP-PIXEL methods included for comparison.

Dimer	S22	SAPT	PIXEL
		kJ mol^{-1}	
2-pyridoxine 2−aminopyridine complex	−69.91	−71.42	−64.8
Adenine thymine complex stack	−51.17	−53.97	−37.2
Adenine thymine Watson-Crick complex	−68.49	−69.33	−72.4
Ammonia dimer	−13.26	−12.22	−17.6
Benzene - Methane complex	−6.28	−5.82	−4.8
Benzene ammonia complex	−9.83	−9.20	−10.2
Benzene dimer parallel displaced	−11.42	−11.97	−4.6
Benzene dimer T-shaped	−11.46	−11.97	−8.9
Benzene HCN complex	−18.66	−18.54	−20.0
Benzene water complex	−13.72	−12.76	−16.8
Ethene dimer	−6.32	−5.40	−7.4
Ethene ethyne complex	−6.40	−5.65	−7.4
Formamide dimer	−66.78	−64.68	−55.1
Formic acid dimer	−77.86	−76.61	−70.9
Indole benzene complex stack	−21.84	−21.71	−9.7
Indole benzene T-shape complex	−23.97	−23.68	−26.8
Methane dimer	−2.22	−2.05	−2.0
Phenol dimer	−29.50	−29.33	−33.3
Pyrazine dimer	−18.49	−18.62	−13.0
Uracil dimer h-bonded	−85.65	−82.97	−79.7
Uracil dimer stack	−41.34	−43.60	−30.9
Water dimer	−21.00	−19.29	−28.5

Some benchmark dimer energies have been collated in databases such as the S22 dataset (Table 1.3), which contains 22 small molecule dimer reference energies.[31,32] To produce these datasets without issues arising from the BSSE, monomer and dimer energies were calculated using basis sets that satisfy convergence to the theoretical infinite limit, known as the Complete Basis Set (CBS) limit. This limit is determined by extrapolating data produced by running several calculations over a range of basis-sets of increasing size (Figure 1.15).[33]

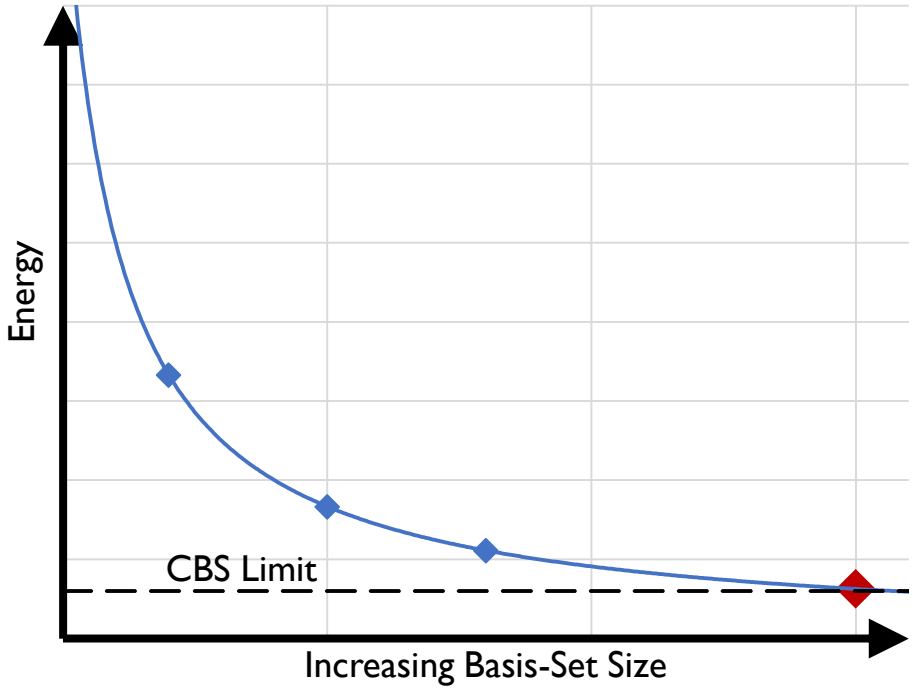


Figure 1.15: Extrapolation of reference basis-set values (blue markers) to find the basis-set necessary (red marker) with satisfactory tolerance of CBS limit (black dashed line).

This approach becomes very computationally expensive for routine calculations or for calculations involving larger molecules. In practice, a single methodology, which is known from the literature or previous work to produce accurate results for similar systems, is applied.

While datasets such as the S22 database provide accurate benchmark and reference energies, generating a new series of dimer energies using the same approach used to form this dataset is not an ideal process. As already discussed, the computational time involved is prohibitive for larger molecules and an increasing number of dimers. Additionally, the energy produced is not broken down into the constituent terms (electrostatic, polarisation, dispersion, and repulsion) and therefore gives very little insight into the nature of the interaction studied.

More intensive ab initio methods have been developed that give an accurate breakdown of energy terms, with perhaps the most successful being Symmetry-Adapted Perturbation Theory (SAPT).

1.4.2. Symmetry Adapted Perturbation Theory

Symmetry-Adapted Perturbation Theory (SAPT) provides an energy breakdown into the four terms discussed in Section 1.1. The sum of these terms forms the total dimer interaction energy. The theory behind SAPT is often noted for its complexity, and a detailed description has been produced by Jeziorski *et al.*[34-36] SAPT is performed by perturbating the wavefunctions of each monomer with respect to each other. The separate energy terms are obtained from the second order perturbation (known as SAPT0). Higher SAPT calculations are possible, in the form of SAPT2, SAPT2+, and SAPT2+3, which incorporate a greater level of perturbation which adds higher order components, such as second order intramonomer electron correlation.[37,38] Higher levels of SAPT are necessary for accurate results when dealing with structures containing large electrostatic terms (e.g. dimers with hydrogen bonding), though the computational cost increases rapidly with the size of the molecules forming the dimer.[38]

As an advantage over the quantum mechanical approaches described previously, which calculate the interaction energy from separate monomer and dimer energies using Equation 1.11, SAPT determines the interaction energy through the perturbation of the entire system, removing the possibility of BSSE.[39]

SAPT provides nearly unrivalled accuracy in the determination of dimer energies, matching some of the highest levels of theory feasible using the traditional quantum mechanical approaches discussed in Section 1.4.1 at a fraction of the computational cost (Figure 1.16).

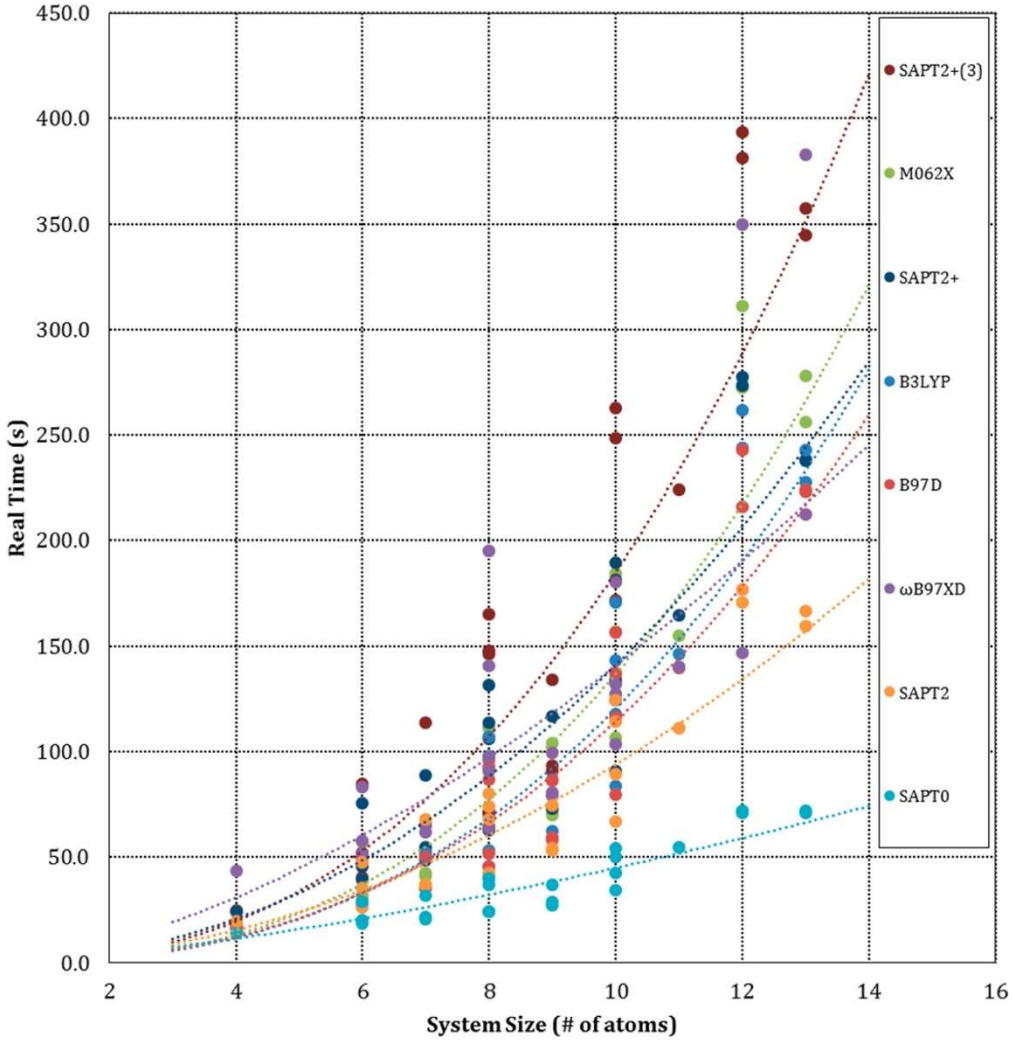


Figure 1.16: Computational time of various computational methods with increasing size of system as produced by Li *et. al.*[40]

Though the improvement in computational time over high level DFT and MP methods is considerable, as with the previous methods, SAPT computation times on current hardware become substantial with larger molecules. This makes the calculation of a full lattice energy (i.e. all dimers out to a convergence point, where the contribution of each individual dimer is effectively 0), prohibitive for most structures. As a result, semi-empirical methods provide a faster, computationally lighter process at the expense of an acceptable decrease in accuracy.

1.4.3. The CLP-PIXEL Method for Structural Analysis

CLP-PIXEL developed by Gavezzotti, is a suite of software for the semi-empirical evaluation of intermolecular interactions in solid-state structures.[8] Like SAPT, PIXEL breaks interaction energies down into components (electrostatic, polarisation, dispersion, and repulsion). Results include lattice energies that accurately reproduce enthalpies found in organic and metal-organic solids (Figure 1.17) at a fraction of the computation cost needed for higher level calculations such as SAPT and DFT.[41]

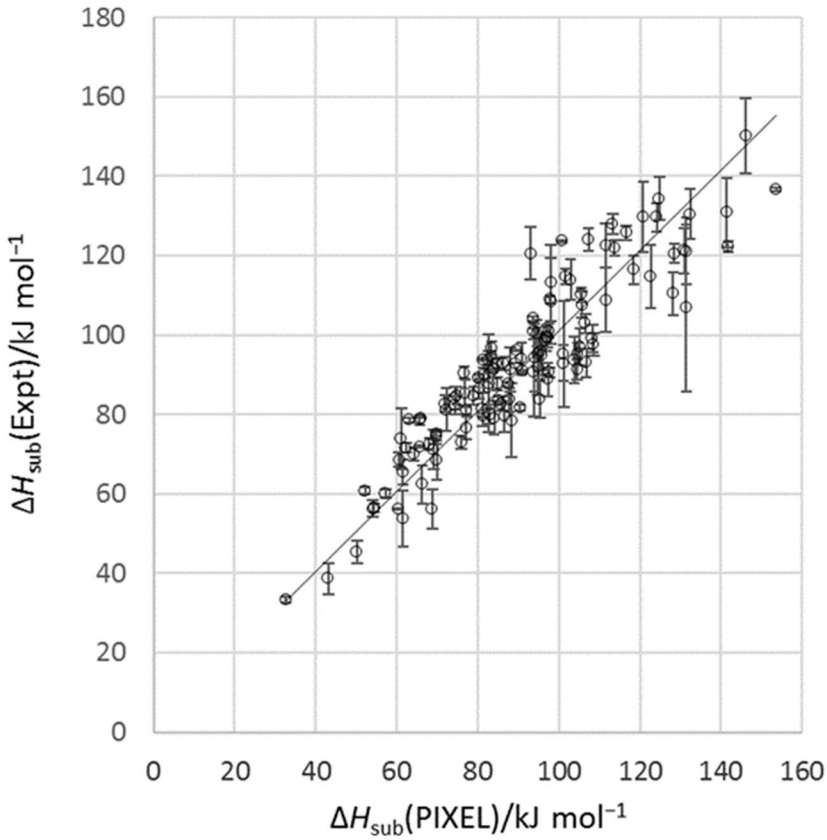


Figure 1.17: Comparison of experimental sublimation energies for a range of organic compounds compared to values obtained by PIXEL. The trend line has equation ΔH_{sub}(Expt) = 1.0118 ΔH_{sub}(PIXEL), and the correlation coefficient is 0.90.[42]

PIXEL calculations are carried out using electron density values of each isolated molecule in the gas-phase usually calculated at the MP2/6-31G** or B3-LYP/6-31G** levels of theory.[20] The results are a 3 dimensional grid of electron density cubes with a standard size of 0.08 Å^3 (default step-size). To reduce the computational load, these pixels are clustered into n x n x n "Super-Pixels", where n is known as the condensation level and has a typical value of 4, so as to produce a grid of super-pixel

electron densities, each with dimensions 0.32 Å. The accuracy of PIXEL results may be improved by decreasing the condensation level, effectively increasing the resolution. The trade-off is an exponentially increasing computational time with decreasing condensation value, *n*. Super-pixels with charge densities below a threshold (10^{-6} electrons) are set to 0, with their electron density being distributed between remaining super-pixels. This reduces the number of pixel-pixel calculations needed. The density map is transformed using the crystal symmetry operations to produce a cluster of energy density out to a user-defined radius (typically ≈ 18 Å for small organic molecules). Calculations are carried out on each combination of super-pixel pairs between molecules (Figure 1.18).

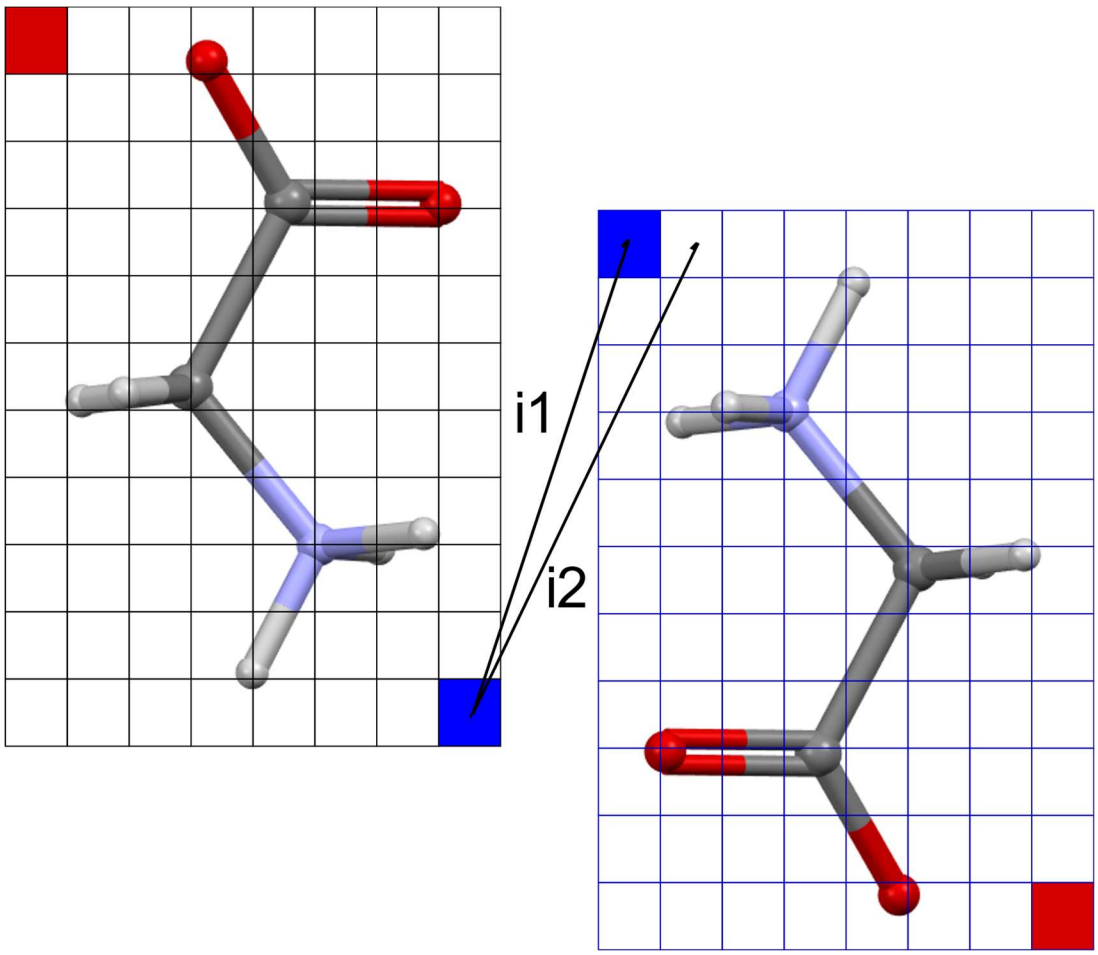


Figure 1.18: Illustrative process of PIXEL calculations on a dimer of α-glycine. Electron density grids are overlaid on reference molecule (black) and dimer molecule (blue fill). With first pixel (blue) and last pixel (red fill) highlighted to show transformation of initial density grid to the dimer molecule. First two pixel-pixel calculations (i1 and i2) are displayed.

Each component is calculated separately using the fundamental energy calculations as laid out below.

The electrostatic term, is formed as the sum of pixel-pixel, pixel-nucleus, and nucleus-nucleus calculations using the Coulomb's law described in Section 1.1.1 (Equation 1.1). Note that for PIXEL, q_1 and q_2 refer to the respective charges of each super-pixel, and r is the separation distance between the centre of each super-pixel. In cases where monomer pixel grids overlap, producing an infeasibly short distance r, the distance is reset to half the pixel-step size in what is known as the "collision-avoidance procedure".

The polarizability term is determined by assigning each pixel, i, to an atom in the molecule, as determined by the atom that gives the smallest ratio of distance to atomic radius. The corresponding polarizability for the pixel is approximated using the following equation:

$$\alpha_i = \left(\frac{q_1}{Z_{atom}}\right)\alpha_{atom} \qquad 1.12$$

Where α_{atom} is the polarizability of the atom as defined by PIXEL, Z_{atom} is the charge on the atom, and q_1 is the charge of the pixel. From this, a polarisation term for the pixel can be determined using the total electric field exerted by surrounding molecules, ε_i, in the equation:

$$E_{pol} = -\frac{1}{2}\alpha_i \varepsilon_i^2 \qquad 1.13$$

Like the electrostatic term, the polarisation term is subject to the collision avoidance procedure which removes instances where overlapping pixels result in interaction energies that tend to infinity. In addition to this, for the polarisation term there is a dampening criterion which compensates for the effective shift in electron density caused by the pixel-pixel distance reset.

The dispersion term is calculated as the summation of pixel-pixel (A_i,B_j) components:

$$E_{Disp(A_iB_j)} = E_{OS} f(R) - \frac{3}{4}\frac{\alpha_i \alpha_j}{(4\pi\varepsilon^0)^2 (R_{ij})^6} \qquad 1.14$$

Where E_{OS} is the oscillator strength as determined from the 'formal' ionisation potential, I_i in Equation 1.15. In order to compensate for the drop in ionisation

potential with increasing distance from the atomic nucleus, the formal ionisation potential is calculated from the ionisation potential of the free atom, I^0, the atomic nucleus-pixel distance, R_i, and the fitted parameter β 1.16. The final component, *f(R)* is defined using 1.17 and adjustable empirical parameter, *D*.

$$E_{OS} = (I_i I_j)^{1/2} \quad 1.15$$

$$I_i = I^o \exp(-\beta R_i) \quad 1.16$$

$$f(R) = \exp -\left(\frac{D}{R_{ij} - 1}\right)^2 \quad 1.17$$

The molecule-molecule dispersion can now be calculated as the sum of pixel-pixel dispersion values.

The final term, repulsion, is calculated from the overlap of charge densities, and equates as:

$$S_{AB} = \sum_{i,A} \sum_{j,B} \Big(\rho_i(A)\rho_j(B)\Big) V \quad 1.18$$

Where ρ_i and ρ_j are the electron charge densities for molecules *A* and *B* respectively, and *V* is the pixel volume. In contrast to the other terms, the repulsion component is calculated using the uncondensed electron density results from Gaussian and is thus unaffected by condensation level but strongly influenced by electron density step-size. As with the polarisation term, each pixel is assigned to a nucleus in the molecule, and so the repulsion is determined as its contribution from atom pairs m and n for molecules *A* and *B* respectively:

$$E_{Rep(mn)} = (K_1 - K_2 \Delta x_{mn}) S_{mn} \quad 1.19$$

$$E_{Rep(AB)} = \sum_{mn} E_{Rep(mn)} \quad 1.20$$

where K_1 and K_2 are positive disposable parameters, and $\Delta\chi_{mn}$ is the difference in Pauling electronegativity. The total interaction Energy, E_{tot} can now be equated as the summation of the electrostatic, polarisation, dispersion, and repulsion terms 1.21. The summation of interaction energies out to the cluster radius gives the lattice energy E_{Lat}.

$$E_{Tot} = E_{Elec} + E_{Pol} + E_{Disp} + E_{Rep} \quad 1.21$$

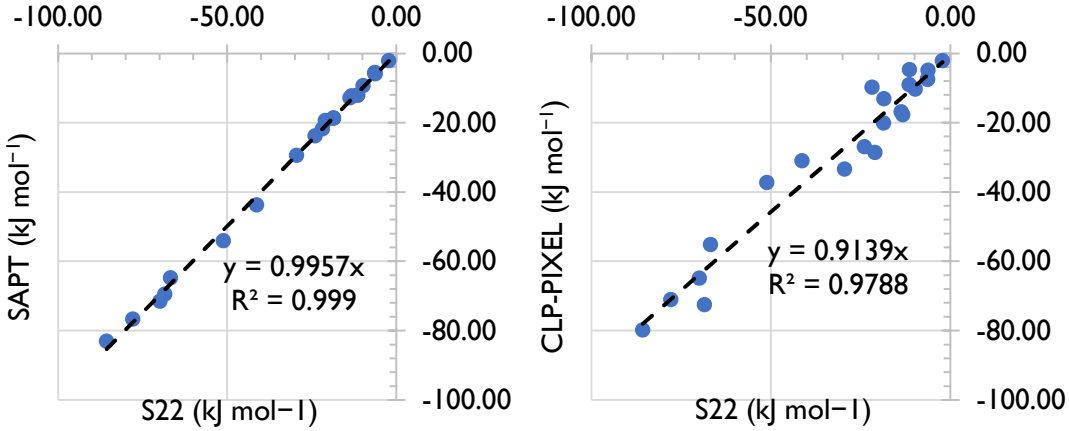


Figure 1.19: Comparison of S22 dataset values with SAPT2+3 (left) and CLP-PIXEL (right). All PIXEL calculations were based on MP2/6-31G** electron densities and condensation level 4.

While PIXEL does not reproduce values from the S22 dataset as closely as SAPT (Figure 1.19), it produces results at a fraction of the computational cost, with lattice energy values within chemical accuracy (Figure 1.17). Moreover, the relative strengths of dimer energies are consistent with the higher level calculations, and the breakdown of individual energy terms, while not matching exactly with SAPT, agree about the physical nature of an interaction.

1.4.4. Interaction energies from CrystalExplorer

As with CLP-PIXEL, CrystalExplorer facilitates the calculation of intermolecular interaction energies in solid state structures with a breakdown of energies into the same components (E_{Elec}, E_{Pol}, E_{Disp}, and E_{Rep}). CrystalExplorer also makes use of ab initio processes such as MP2 and DFT, either through Gaussian or Tonto programs.

The electrostatic term is calculated to the classical approach of monomer charge interactions, similar to PIXEL. The same is true of the polarisation energy, which is estimated over the sum of all atom-atom interactions, as determined using Equation 1.13. The dispersion term is generated from Grimme's D2 correction summed over all atom-atom interactions.[43] The final term, repulsion, comes from Su & Li's method using the antisymmetric product of monomer spin orbitals.[44]

The relative sizes of each term, and the summed total interaction energy do not relate well to higher level calculations such as SAPT, and therefore scaling factors (k_{term}) are added to each of the terms which are fitted from comparisons between the above process and higher level calculations for 232 molecular dimers.[45-47] The interaction energy is thus produced using Equation 1.22.

$$E_{Tot} = k_{Elec}E_{Elec} + k_{Pol}E_{Pol} + k_{Disp}E_{Disp} + k_{Rep}E_{Rep} \qquad 1.22$$

The interaction energies and individual breakdown of energy terms closely resemble those from PIXEL-CLP, with a similar level of accuracy and computational expense. These results can be displayed pictorially using energy frameworks where the relative strengths of interactions in the crystal structure are clearly visible (Figure 1.20).[3] Frameworks can also be generated for each of the separate terms. This can be used to determine the first coordination sphere from repulsive terms for example, or in quickly assessing the dominating terms across the crystal structure.

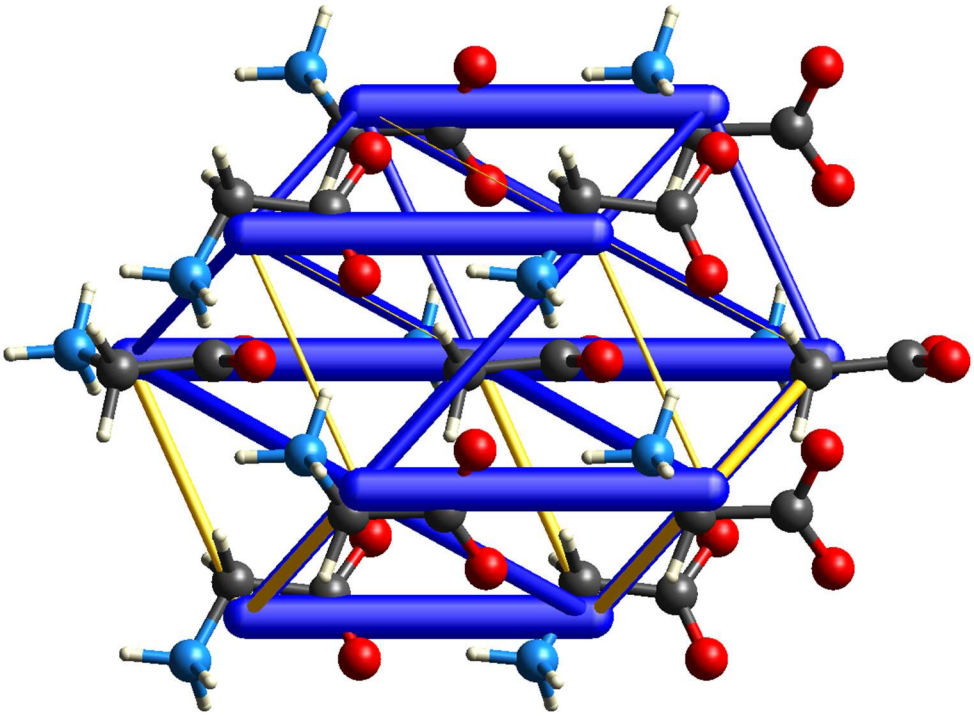


Figure 1.20: Energy framework for γ-glycine as displayed by CrystalExplorer. Strut thickness represents relative molecule-molecule interaction strengths. Blue struts represent stabilizing (< 0 kJ mol^{-1}) interactions, where yellow struts represent destabilizing interactions (> 0 kJ mol^{-1}).

1.5. Thesis Aims

The previous section has demonstrated many approaches that may be taken to study intermolecular interactions in solid-state materials. It has emphasised the value of computational techniques which allow the user to describe interactions both in terms of nature and quantifiable strengths.

The use of computational techniques for organic structures is already well established, with software such as PIXEL and CrystalExplorer being widely used in the rationalization and description of their crystal packing and properties. Use of these methods on transition metal containing compounds is less common, in part due to the difficult of searching for oxidation state spesific structures as described in Section 1.1. In addition, while CrystalExplorer provides a simply interface and process for setting up computaional calculations, it restricts the user to using local computational resources, which become prohibitivly slow when looking at larger systems such as co-ordination complexes. PIXEL, on the other hand, facilitates the passing of computational tasks to external clusters with greater computational power, but has no intuitive interface, which increases calculation set-up time due to the need to manually create input files, and introduces the potential for significant user-error.

This thesis aims to improve the accesability of CSD based structural analysis, and demonstrate the spesific value of this approach to coordination compounds. To achieve these aims, this thesis will first aim to improve the situation regarding missing oxidation states in the CSD, incoperating techniques such as BVS to develop an automated procedure for oxidation state determination.

To further ease the analysis of crystal structures, the second task is to develop a new interface and process for the automation of PIXEL calculations within the CSD software Mercury.

Lastly, and in order to facilitate the second aim of demonstrating the value of intermolecular interaction analysis towards coordination compouds, PIXEL will be used to investigate and rationalize the role of intermolecular interactons in the behavior of spin-crossover matierals.

1.6. References

(1) Kennard, O. *From Private Data to Public Knowledge In: Butterworth, I ed. The Impact of Electronic Publishing on the Academic Community.*; Portland Press Ltd., **1997**.

(2) Kvick, A.; Canning, W. M.; Koetzle, T. F.; Williams, G. J. B. An experimental study of the influence of temperature on a hydrogen-bonded system: the crystal structure of [gamma]-glycine at 83 K and 298 K by neutron diffraction. *Acta Crystallographica Section B* **1980,** *36* (1), 115.

(3) Mackenzie, C. F.; Spackman, P. R.; Jayatilaka, D.; Spackman, M. A. CrystalExplorer model energies and energy frameworks: extension to metal coordination compounds, organic salts, solvates and open-shell systems. *IUCrJ* **2017,** *4* (5), 575.

(4) Bacon, G. E.; Curry, N. A.; Wilson, S. A.; Spence, R. A crystallographic study of solid benzene by neutron diffraction. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **1964,** *279* (1376), 98.

(5) Langan, P.; Mason, S. A.; Myles, D.; Schoenborn, B. P. Structural characterization of crystals of [alpha]-glycine during anomalous electrical behaviour. *Acta Crystallographica Section B* **2002,** *58* (4), 728.

(6) Dawson, A.; Allan, D. R.; Belmonte, S. A.; Clark, S. J.; David, W. I. F.; McGregor, P. A.; Parsons, S.; Pulham, C. R.; Sawyer, L. Effect of High Pressure on the Crystal Structures of Polymorphs of Glycine. *Crystal Growth & Design* **2005,** *5* (4), 1415.

(7) Jones, J. E.; Chapman, S. On the determination of molecular fields. —II. From the equation of state of a gas. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **1924,** *106* (738), 463.

(8) Angelo, G. Calculation of lattice energies of organic crystals: the PIXEL integration method in comparison with more traditional methods. *Zeitschrift für Kristallographie - Crystalline Materials* **2005,** *220* (5-6), 499.

(9) Thakur, T. S.; Dubey, R.; Desiraju, G. R. Intermolecular atom-atom bonds in crystals - a chemical perspectiveThis essay and that by Lecomte et al. (2015) comment on Dunitz (2015). *IUCrJ* **2015,** *2* (2), 159.

(10) Bondi, A. van der Waals Volumes and Radii. *The Journal of Physical Chemistry* **1964,** *68* (3), 441.

(11) Alvarez, S. A cartography of the van der Waals territories. *Dalton Trans* **2013,** *42* (24), 8617.

(12) Harding, M. M.; Long, H. A. The crystal and molecular structure of l-cysteine. *Acta Crystallographica Section B* **1968,** *24* (8), 1096.

(13) Nyburg, S. C.; Faerman, C. H. A revision of van der Waals atomic radii for molecular crystals: N, O, F, S, Cl, Se, Br and I bonded to carbon. *Acta Crystallographica Section B* **1985,** *41* (4), 274.

(14) Pauling, L. THE PRINCIPLES DETERMINING THE STRUCTURE OF COMPLEX IONIC CRYSTALS. *Journal of the American Chemical Society* **1929,** *51* (4), 1010.

(15) Brown, I. D. *The Chemical Bond in Inorganic Chemistry: The Bond Valence Model*; Oxford University Press, 2016.

(16) Keith, G. M.; Kirk, C. A.; Sarma, K.; Alford, N. M.; Cussen, E. J.; Rosseinsky, M. J.; Sinclair, D. C. Synthesis, Crystal Structure, and Characterization of

Ba(Ti1/2Mn1/2)O3: A High Permittivity 12R-Type Hexagonal Perovskite. *Chemistry of Materials* **2004,** *16* (10), 2007.

(17) Brown, I. D. Recent Developments in the Methods and Applications of the Bond Valence Model. *Chemical Reviews* **2009,** *109* (12), 6858.

(18) Brown, I. D. *The Chemical Bond in Inorganic Chemistry: The Bond Valence Model*; Oxford University Press, 2006.

(19) Brown, I. D. IUCR, 2016; Vol. 2019.

(20) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.et al. Wallingford, CT, 2016.

(21) Jayatilaka, D.; Grimwood, D. *Tonto: A Fortran Based Object-Oriented System for Quantum Chemistry and Crystallography*, 2003.

(22) Kristyán, S.; Pulay, P. Can (semi)local density functional theory account for the London dispersion forces? *Chemical Physics Letters* **1994,** *229* (3), 175.

(23) Pérez-Jordá, J.; Becke, A. D. A density-functional study of van der Waals forces: rare gas diatomics. *Chemical Physics Letters* **1995,** *233* (1), 134.

(24) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *The Journal of Chemical Physics* **2010,** *132* (15), 154104.

(25) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *Journal of Computational Chemistry* **2011,** *32* (7), 1456.

(26) Kovács, A.; Cz. Dobrowolski, J.; Ostrowski, S.; Rode, J. E. Benchmarking density functionals in conjunction with Grimme's dispersion correction for noble gas dimers (Ne2, Ar2, Kr2, Xe2, Rn2). *International Journal of Quantum Chemistry* **2017,** *117* (9), e25358.

(27) Moellmann, J.; Grimme, S. DFT-D3 Study of Some Molecular Crystals. *The Journal of Physical Chemistry C* **2014,** *118* (14), 7615.

(28) Jansen, H. B.; Ros, P. Non-empirical molecular orbital calculations on the protonation of carbon monoxide. *Chemical Physics Letters* **1969,** *3* (3), 140.

(29) Boys, S. F.; Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Molecular Physics* **1970,** *19* (4), 553.

(30) Meunier, A.; Lévy, B.; Berthier, G. Corrélation électronique et effets de base dans l'étude de la liaison hydrogène: le dimère mixte ammoniac-eau. *Theoretica chimica acta* **1973,** *29* (1), 49.

(31) Rezac, J.; Jurecka, P.; Riley, K.; Černý, J.; Valdés, H.; Pluháčková, K.; Berka, K.; Řezáč, T.; Pitoňák, M.; Vondrasek, J.et al. Quantum Chemical Benchmark Energy and Geometry Database for Molecular Clusters and Complex Molecular Systems (www.begdb.com): A Users Manual and Examples. *Collection of Czechoslovak Chemical Communications* **2008,** *73*, 1261.

(32) Jurečka, P.; Šponer, J.; Černý, J.; Hobza, P. Benchmark database of accurate (MP2 and CCSD(T) complete basis set limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs. *Physical Chemistry Chemical Physics* **2006,** *8* (17), 1985.

(33) Vasilyev, V. Online complete basis set limit extrapolation calculator. *Computational and Theoretical Chemistry* **2017,** *1115*, 1.

(34) Stone, A. J. The theory of intermolecular forces. **2013**.

(35) *SAPT: a program for many-body symmetry-adapted perturbation theory calculations of intermolecular interaction energies*, 1993.

(36) Perturbation theory approach to intermolecular potential energy surfaces of Van der Waals complexes. *Chemical reviews.* **1994,** *94*, 1887.

(37) Hohenstein, E. G.; Sherrill, C. D. Efficient evaluation of triple excitations in symmetry-adapted perturbation theory via second-order Møller–Plesset perturbation theory natural orbitals. *The Journal of Chemical Physics* **2010,** *133* (10), 104107.

(38) Parker, T. M.; Burns, L. A.; Parrish, R. M.; Ryno, A. G.; Sherrill, C. D. Levels of symmetry adapted perturbation theory (SAPT). I. Efficiency and performance for interaction energies. *The Journal of Chemical Physics* **2014,** *140* (9), 094106.

(39) Szalewicz, K. Symmetry-adapted perturbation theory of intermolecular forces. *WIREs Computational Molecular Science* **2012,** *2* (2), 254.

(40) Li, A.; Muddana, H. S.; Gilson, M. K. Quantum Mechanical Calculation of Noncovalent Interactions: A Large-Scale Evaluation of PMx, DFT, and SAPT Approaches. *Journal of Chemical Theory and Computation* **2014,** *10* (4), 1563.

(41) Chickos, J. S.; Gavezzotti, A. Sublimation Enthalpies of Organic Compounds: A Very Large Database with a Match to Crystal Structure Determinations and a Comparison with Lattice Energies. *Crystal Growth & Design* **2019,** *19* (11), 6566.

(42) van Gerwen, P. MChem Report. **2017**.

(43) Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *Journal of Computational Chemistry* **2006,** *27* (15), 1787.

(44) Su, P.; Li, H. Energy decomposition analysis of covalent bonds and intermolecular interactions. *The Journal of Chemical Physics* **2009,** *131* (1), 014102.

(45) Turner, M. J.; Grabowsky, S.; Jayatilaka, D.; Spackman, M. A. Accurate and Efficient Model Energies for Exploring Intermolecular Interactions in Molecular Crystals. *The Journal of Physical Chemistry Letters* **2014,** *5* (24), 4249.

(46) Acree, W.; Chickos, J. S. Phase Transition Enthalpy Measurements of Organic and Organometallic Compounds. Sublimation, Vaporization and Fusion Enthalpies From 1880 to 2015. Part 1. C1 − C10. *Journal of Physical and Chemical Reference Data* **2016,** *45* (3), 033101.

(47) Acree, W.; Chickos, J. S. Phase Transition Enthalpy Measurements of Organic and Organometallic Compounds and Ionic Liquids. Sublimation, Vaporization, and Fusion Enthalpies from 1880 to 2015. Part 2. C11–C192. *Journal of Physical and Chemical Reference Data* **2017,** *46* (1), 013104.

2. Automated Oxidation-State Assignment for Metal Sites in Coordination Complexes in the Cambridge Structural Database

2.1. Introduction

The Cambridge Structural Database (CSD)[1] currently contains over 400,000 structures of coordination complexes, but only about half of these entries specify the metal oxidation state, and this is reported in the compound name field. Although the current system of incorporating oxidation states in the entry compound name provides some scope for filtering entries, it would be much more advantageous to associate specific oxidation states with individual transition metal sites. In this way it would be possible to distinguish specific oxidation states in polynuclear complexes as well as quickly filter entries by both metal and associated valence.

As the CSD moves to a new data structure which includes oxidation state as an atomic property, new processes are needed to generate and assign individual valences. Given the number of transition metal-containing entries, it would be impractical to attempt manual identification and curation of current, as well as future, entries. Automated processes that can distinguish individual atomic valences are therefore highly desirable.

Past studies have identified and validated transition metal oxidation states using a combination of Bond Valence Sum (BVS) and ligand templating processes, implementing a two-step process in oxidation state assignment and validation.[2] An initial oxidation state was estimated using a ligand templating method whereby an algorithm was trained to recognise common ligand templates surrounding a metal centre, and then apply the charge associated with each ligand to determine the corresponding charge of the metal. Having achieved this, BVS was applied to the structure using parameters associated with the oxidation state interpreted from the results of the ligand template method.

This method was applied to a subset of 743 manually-verified copper (+1) and (+2) structures, with 98% successful assignment. While these results are extremely promising, the procedure relied on appropriate coverage of ligand templates and BVS parameters to produce a confident assignment. Where either the template or BVS method deviated from the expected value, manual inspection was required to check the assignments made.

The BVS method has also been applied to inorganic compounds in the ICSD with the aim of predicting the formation of likely oxidation states in the presence specific anions [3].

Presented here are new methods that can be broadly applied to molecular coordination complexes, in most cases without the need for manual intervention. As in previous work, oxidation states are assigned using the BVS method, but without the need to assume or derive an initial oxidation state estimate. The BVS calculations are supported by assignment of ligand charges, but avoid the definition of templates (in most cases, open-shell ligands are an exception). The combination of these methods provides a confidence-scored oxidation state for each metal atom present in a complex. All calculations make use of the CSD Python API (application programming interface), which has been distributed alongside the CSD since 2015.

2.2. Methodology

2.2.1. Assignment of oxidation states using the bond valence method

In the bond valence method each metal(i)-ligand(j) bond is assigned a valence, S_{ij}, based on its length and two empirical parameters, R_0 and B. The sum of the valences of the bonds formed by the metal is its oxidation state [4]. Bond valence parameters depend on the metal, its oxidation state and the identity of the bonded ligand atom.

The parameters R_0 and B are taken from the database compiled by I. D. Brown,[5] and bonds are defined using the default CSD chemical connectivity cut-offs. The calculation is carried out for all available bond valence parameters provided that, for each metal-ligand bond in the molecule, parameters exist for all common oxidation states. An oxidation state was considered 'common' if it applied to more than 15% of a metal's assigned entries in the CSD (a list is given in the SI (Table S2.1); this choice of cut-off gives a listing broadly similar to that in Housecroft and Sharpe's popular textbook [6]). The value of S_{ij} is calculated for each of the oxidation states for which values of R_0 and B are available in Brown's database (Equation 2.1, where R_{ij} is the metal-ligand bond distance).

$$S_{ij} = e^{\frac{(R_0 - R_{ij})}{B}} \qquad 2.1$$

For example, for the chromium compound HIQYAJ contains the $[Cr(oxalate)_2(H_2O)_2]^-$ anion.[7] The common oxidation states for Cr are +2 and +3, and so unless parameters for Cr-O bonds for both are available no attempt is made to assign the oxidation state at all. In fact parameters are available for Cr-O bonds in all oxidation states from Cr(+2) to Cr(+6), and all of these are considered in the oxidation state assignment procedure.

For each oxidation state, the values of S_{ij} are summed to give a total trial oxidation state, V_t. The value of *BVS* is compared to the oxidation state (V_p) corresponding to the bond valence parameters used to calculate them.

$$\Delta = \left| V_t - V_p \right| \qquad 2.2$$

The oxidation state of the metal is taken as the one which yields the smallest value of Δ, that is, the oxidation state which is most consistent with the parameters used to calculate it (Equation 2.2).

For example, in the four co-ordinate cobalt complex KUYHES [8] there are two Co-N bonds with distances 2.042 and 2.053 Å and two Co-Cl bonds measuring 2.219 and 2.217 Å. Cobalt has two common oxidation states, +2 and +3, and Co-N and Co-Cl bond valence parameters are available for both. Using the Co(+2)-N and Co(+2)-Cl bond valence parameters to calculate the bond valences of the Co-N and Co-Cl bonds yields a total trial valence (V_t) of 1.987. The difference, Δ, between this and the reference oxidation state used to select the bond valence parameters (V_p) is $|1.987 - 2| = 0.013$. Using the bond valence parameters for Co(+3)-N and Co(+3)-Cl bonds yields $V_t = 2.014$ and $\Delta = |2.014 - 3| = 0.986$. Since $0.013 < 0.986$, the oxidation state of the cobalt is taken as +2.

If the minimum value of Δ is greater than 0.5 a warning is added to the assignment. Warnings are used in confidence scoring (see Section 2.2.3).

For many bond types, R_0 and B have been determined multiple times. Different parameters may apply to different spin states, *e.g.* high and low spin Fe(+2)-N, or be derived from different classes of compound or data-sets of different sizes. Each available set of parameters was used to calculate a value of Δ, with the smallest value being used to assign the oxidation state of the metal. The procedure, which was found to reproduce known oxidation states more reliably than using the parameters designated "most reliable" in Brown's database, is exemplified in detail in the SI for the KUYHES example of the previous paragraph (Table S2.2, S2.3).

2.2.2. Assignment of oxidation states using ligand charges

Ligand charge assignment procedure

As an alternative to the bond valence method, likely ligand charges were also determined using the very fast semi-empirical electronic structure package MOPAC [9]. The overall charge on a complex can be derived from the sum of the formal atomic charges that are stored in the CSD as *atomic charge properties*. Therefore, the metal

oxidation state can be assigned as (the sum of the stored formal atomic charges) – (sum of the ligand charges).

In the first stage of the procedure the metal centre is removed, leaving the ligand fragments for charge assignment. This process only considers ligands directly connected to the metal centre of interest. Where a salt occurs in the database, formal atomic charges are added to the metal centre by the scientific editors at the CCDC to achieve a charge-neutral structure.

For each ligand fragment the total number of electrons is determined assuming charge-neutrality. For ligands with an even number of electrons, possible charges were taken to be +4 to −8 in steps of 2e; for those with an odd number of electrons possible charges were +3 to −9, also in steps of 2e. This procedure does not consider the possibility that a ligand has an open-shell (*i.e.* a radical) electron configuration. Radicals are discussed below, along with further comments on cationic ligands. For each charge, a single-point electronic structure calculation (MOPAC) is carried out using the crystal structure geometry and the PM7 method [10]. Each calculation yields a heat of formation and a Parr and Pople hardness parameter as well as the energies of the highest occupied and lowest unoccupied molecular orbitals. The charge is assigned on the basis of the formation energy and the hardness parameter.

Parr & Pople hardness charge assignment

The hardness parameter quantifies the resistance to changes in the electron configuration [11], and the charge on the ligand was taken as the one yielding the largest hardness parameter. Hardness values typically fall into the range 0-10 eV. Any charges yielding a hardness outside this range are disregarded. A warning is issued if the difference in hardness is less than 1eV.

Charge assignment using formation energy

The charge on the ligand was taken as the one yielding the most negative formation energy. As an example, the energy versus charge plot for a ligand with formula 'NO_3' is shown in Figure 2.1. There is a clear minimum for a charge of −1, indicating that the ligand is NO_3^-.

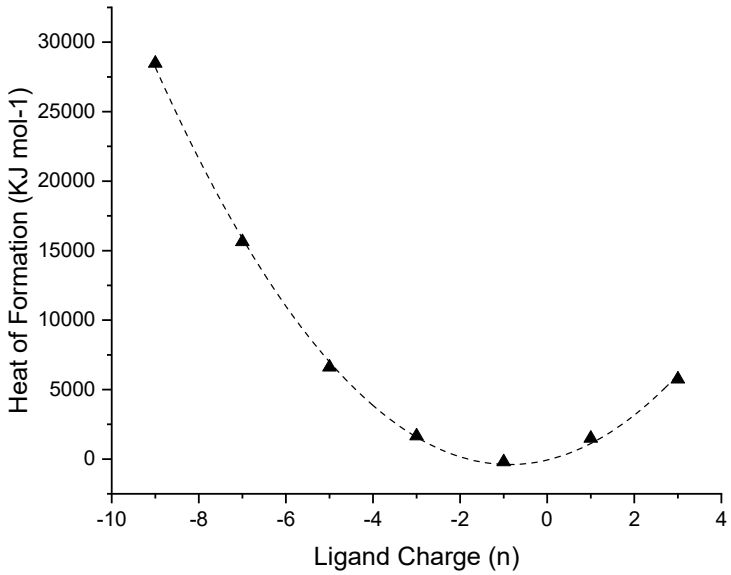


Figure 2.1: Heat for formation (kJ mol^{-1}) versus ligand charge (*n*) for NO_3^n. Energies calculated in single-point energy calculations in the crystal structure geometry using MOPAC.

For some structures, a ligand fragment may produce a set of formation energies with a shallow minimum, making charge assignment ambiguous. Experience showed that ambiguities arose when the energy difference between charges was lower than 150 kJ mol^{-1}. Values lower than this could, for example, lead to assignment of different charges for identical ligands in different crystal structures. Where energy differences do suggest a shallow minimum, a warning is added to the fragment assignment, and this is carried forward when considering overall assignment confidence.

Assignment of oxidation states using hydrogen placement algorithms

The CSD Python API has a built-in molecular editing tool for automatic hydrogen placement which can be used to determine the charge of the ligand following removal of the metal atoms (as in Section 2.2.1). The number of H-atoms in a ligand is first recorded. The H-atoms are all removed and then replaced using the H-atom generation routine assuming charge neutrality. The difference between the number of H-atoms before and after this procedure is the charge. For example, the methoxide ligand CH_3O^- contains three H-atoms. Removal of these followed by automatic H-atom placement generates methanol, CH_3OH, containing four H-atoms. The charge on the original CH_3O fragment is therefore −1 since addition of one

proton is required to generate a neutral molecule. Having determined the ligand charges in this way the oxidation state of the metal is assigned following the procedure of the previous sections.

Radicals

The ligand charge calculation is carried out in steps of 2e because the problems associated with shallow minima become much more common if charges are sampled in steps of 1e. The number of structures containing radical ligands is quite small, <2% of structures in the test-set used in this work for method validation (Section 3.1). For common radical species, these ligands can be identified beforehand from their SMILES formulae and are added manually to an SQLite database in the form of an "exceptions list" look-up table, which pre-assigns a ligand fragment charge before any determination processes are carried out. This procedure is similar to the templating method used by Shields *et al.* in their work.[12]

Cationic ligands

Cations can be readily identified in entries from the CSD by the systematic presence of positive symbols in the SMILES formulae generated by the CSD Python API, so that the charge of the fragment is determined by simply summing the number of positive symbols and subtracting the number of negative symbols within the SMILES formula. As SMILES-based charge labelling requires specifically ionic atomic sites, this process cannot be used to distinguish between neutral and anionic ligands, where the metal-ligand bond is typically considered as neutral.

Where a ligand is zwitterionic, there is an ambiguity as to how ligands have been labelled. In the Zn^{2+} complex CSD refcode EGAPOR,[13] zwitterionic N-n-propylsalicylaldimine-O ligands are identified in the CSD entry, with ligating atoms denoted with a negative charge and the protonated imine nitrogen atom with a positive charge. By contrast, in the Cu(+2)-containing refcode CICWIU,[14] only the positive charge on a terminal ammonium moiety is identified in the SMILES formula "N(=C\c1ccccc1O)/CCNCCNCC[NH3+]", suggests a +1 cation, rather than the true neutral overall ligand. In order to address these issues, assignments made using this method are only accepted providing the closed shell requirement described

earlier is still obeyed. Where this is not the case, a warning is displayed, and the corresponding assignment of metal oxidation state is aborted.

Oxidation state assignment based on ligand charges.

The metal oxidation state is determined for mononuclear complexes from the total charge of the ligands and the overall charge on the complex to achieve a net neutral crystal structure.

The same approach can be applied to polynuclear complexes where there is a single metal atom in the asymmetric unit and assuming charge-order so that the overall charge is split evenly between each metal atom present in the overall structure. As an example, the dimeric Cu complex SAVRIQ01 is located on an inversion centre so that the asymmetric unit contains a single copper atom.[15] Assigning all ligand charges in the complex gives an overall charge of +4. Using the assumption that each asymmetric unit has the same valence, the valence of each copper atom is equal to ½ (*i.e.* 1/n asymmetric units that make up a complete molecule) times the overall charge = +2.

In other polynuclear complexes the total ligand charge can only be used to obtain the sum of the metal oxidation states, and the BVS is the only method capable of assigning the oxidation states of individual metal atoms. The total ligand charge is used instead for validating the BVS assignments.

2.2.3. Oxidation state assignment and confidence scoring

The preceding sections have described four methods for oxidation state assignment: a BVS approach and three ligand charge-based methods using minimum energy, maximum hardness and the number of H-atoms. It is only strictly necessary to apply these methods to a new CSD entry in cases where an author-supplied oxidation state is not available, though it is recommended that they could also be used to validate author-assignments.

Where named valences are not available, assignments are made using a combination of all methods described. In ideal cases, all four methods should agree on the assigned oxidation state. In cases where the methods disagree, the oxidation state assignment is attempted using the BVS method as this is the only method that can be applied to

both the mononuclear and polynuclear complexes. Where BVS cannot be applied, the assignment is made based on the maximum hardness method for ligand charge assignment. The reliability of this method is very similar to the energy assignment method but during testing there were found to be fewer ambiguous cases (see above) than for the energy method.

Following oxidation state assignment, a confidence score is determined based on the success rate of each method, the agreement between different methods, and the occurrence of any warnings. A numeric score is determined using a summation of values from Table 2.1. Each assignment may have an overall score between 0 (no assignment) and 17 (all assignments agree without error). For simplicity, these are banded into letter grades (A-D) as in

Table 2.2, with A indicating the highest level of confidence and D indicating the lowest level of confidence.

The confidence bands have been defined on the basis of experience, based on which methods were most effective at predicting the correct oxidation state. Examples are given below.

Table 2.1: Confidence scoring values for each assignment method. Scores are given for agreement with the most reliable method.

Method	Method Agrees with Assignment (without errors/warnings)
Bond Valence Sum	5(+1)
Mopac Ligand Assignments by Hardness	4(+1)
Mopac Ligand Assignments by Energy	3(+1)
Ligand Assignments by Hydrogen Placement	2

Table 2.2: Confidence score grade bandings. The reliability is based on the results of Section 2.3.1.

Score	Band Values	Description	Reliability
0	U	Unassigned	0%
0-5	D	Very unreliable	18%
6-8	C	Quite unreliable	56%
9-12	B	Reliable	98%
>12	A	Very reliable	>99%

2.2.4. Ligand database

The three charge assignment methods described above have been used to generate an SQLite database of ligands, their frequency in the CSD and the assigned charge. The database can be accessed, updated and added to through the SQLite3 Python Module [16].

The database contains the SMILES formula for each ligand in the CSD and the number of times it has been encountered. The number of entries is currently 12,939. Most ligands appear in multiple CSD entries, yielding a distribution of charges for each of the three methods described in Section 2.2. For each ligand the modes (*i.e.* the most common values) and standard deviations of the each of the three distributions is stored in the database. These data enable a proposed charge assignment to be checked against previous assignments for the same ligand, while also providing a measure of confidence in the comparison.

The database facilitates the ability to over-ride potentially incorrect charge assignments where the value disagrees with previous values. In order to achieve this, for each ligand fragment encountered, the mode of previous assignments is compared to the currently determined value. If the value does not match the most common charge determined in the database, the database value is used instead, with a warning generated, reducing the confidence score by 1.

2.2.5 Confidence scoring examples

Example 1: In the entry AMIRAR [17] where Cu(+1) is coordinated to a acetonitrile and a thiopyridazine scorpionate ligand in which both the N and B atoms are bound to the Cu. The BVS method could not be applied because Cu-B parameters are unavailable. As a result, only the ligand charge methods can be applied. Application of the minimum energy method yields an oxidation state of +3, the hardness method a value of +1 and the H-atom placement method a value of +1; no errors were issued in any of these procedures. The value selected is taken from the hardness method, correctly assigning a valence of +1. The confidence score is the sum of 0 for the BVS method, 5 for the hardness method (4 + 1 for no errors in assignment), 0 for the energy since this disagrees with the results from the method with highest reliability, and 2 for H-placement method. For both ligands there are previous assignments

available in the SQLite database, and for all methods the mode charge agrees with the current assignment. The total score is 7: this C-grade assignment should be considered quite unreliable.

Example 2: In chloro-bis(N-phenylbenzohydroxamato)-(triphenylphosphine)-rhodium(3+) (refcode: CAFSEI [18]), BVS assignment is not possible due to a lack of Rh-P/Rh-Cl parameters and assignment must be made using ligand charge methods only. For the chloride and triphenylphosphine ligands, the hardness method correctly assigns charges of −1 and 0, respectively. However, the N-phenylbenzohydroxamato ligands are incorrectly given a charge of +1. This ligand is listed in the ligand database with 9 previous assignments, with a (correct) mode of −1. The database value charge (−1) replaces that determined by the hardness method, and a warning associated with the hardness method. The hardness confidence score is therefore 5 – 1 = 4. The energy and hydrogen placement methods both yield the correct charge of −1, and so all 3 methods produce the correct valence of Rh(3+) for this structure, the final confidence therefore is a sum of the scores for BVS (0), energy (4), hardness (4) and hydrogen placement (2) = 10, lying in the B confidence band. This B-grade assignment should be considered quite reliable.

2.3. Discussion

2.3.1. Success rate of oxidation state assignment

The aim of the present study was to determine the oxidation states of transition metals in coordination complexes using an automated procedure and to devise a measure of the confidence in the assignments. Compounds containing metal-carbon bonds, nitrosyl ligands or metal-ligand multiple bonds have been excluded, and the methods described apply to 'classic' coordination complexes only, and not to organometallic compounds. The focus on coordination complexes in part simply reflects the lack of bond valence data for organometallic compounds, but oxidation state assignment in organometallic chemistry is also ambiguous: even so fundamental a compound as ferrocene may be considered to contain Fe(0) or Fe(+2).

Two approaches were used in oxidation state assignment, (i) the bond valence method, and (ii) calculation of ligand charges. The bond valence method is applicable to any complex whether it is mono- or poly-nuclear, but it depends on the availability of suitable parameters. Ligand charges were derived using three methods: from the minimum of energy versus charge plots, from Pearson's principle of maximum hardness and from automated hydrogen placement routines. Once the ligand charges are known they can be applied to assign the oxidation state of a metal in a mononuclear complex, but they only yield the total of all the metal oxidation states in polynuclear complexes.

In order to validate and optimise the different methods an initial testing set was generated which contained entries with predefined valences. Suitable entries were extracted from the CSD November 2018 release by scanning for compound names containing a string comprising the name of a transition metal followed by a Roman numeral in parentheses [e.g. 'nickel(ii)']. In addition to entries containing one metal, this approach can be applied to multiple metal structures where more than one valence is specified, provided that only one valence is present for each metal name.

The methods described in this chapter are intended to be applicable to coordination complexes, and entries containing the 'organometallic' moieties listed in the first paragraph of this section were omitted. 3D co-ordinates were required to be present for all atoms, and disordered structures and structures containing errors were

omitted. Entries with missing hydrogen atoms were omitted as well as those where the number of hydrogen atoms present differed from the figure calculated using the CSD structure-editing tools. Where more than one structure is available in a single refcode family, the structure with the lowest *R*-factor was selected.

The final test-set contained 54,999 unique metal environments across 47,716 molecular components, from 43,423 entries. This set contained entries from all transition metals, with a minimum number of 52 environments for scandium and maximum number of 13,259 environments for copper. The test was run twice, first to populate the ligand database with fragment results, and then again to enable the charge validation to be incorporated into the confidence scores.

The overall success rates of each of the four methods for oxidation state assignment are summarised in Table 2.3, where, in order to accommodate both mono- and poly-nuclear complexes, the entries are based on reproduction of the total metal oxidation state. It should be noted that while application of the ligand charge methods was achieved successfully for nearly all entries, bond valence assignments were reliant on the availability of parameters and so has only been applied for 82.50% of components. The ligand charge methods based on hardness and energy are as effective as the traditional BVS approach, and can be applied to all complexes.

Table 2.3: Breakdown of assignment results by method against test-set entries.

Method	Summed Component Valence Assignments			
	Correct	Incorrect	% Correct	Not Applied
BVS	35419	3950	89.97%	8347
Energy	44418	3298	93.09%	0
Hardness	44367	3349	92.98%	0
Hydrogen-Placement	41802	5914	87.6%	0
Overall Assignment	43220	4496	90.58%	0

The applicability of the BVS method varies significantly across the periodic table, with fewer parameters being available for the second and third row transition metals than for the first row (Figure 2.2). The figures in the first row of Table 2.3 are thus weighted towards complexes of the 3d metals. The BVS method is always needed for assignment of individual metal sites in polynuclear complexes, and the applicability of the methods described in this chapter therefore becomes quite patchy for polynuclear complexes of the second and third row metals. Of the complexes in Table 2.3, BVS could not be applied to the metal atoms in 8,347 components. Where these components are polynuclear complexes no assignment can be made at all, meaning that no assignment was made for 3,113 metal sites in the test set (amounting to 5.7% of the set). This situation should improve as bond valence parameters are determined for more element bond types in a range of oxidation states, the results obtained here illustrating the importance of research in this area. The oxidation states of the remaining 8,347 – 3113 = 5,234 mononuclear complexes could all be assigned using the ligand charge methods.

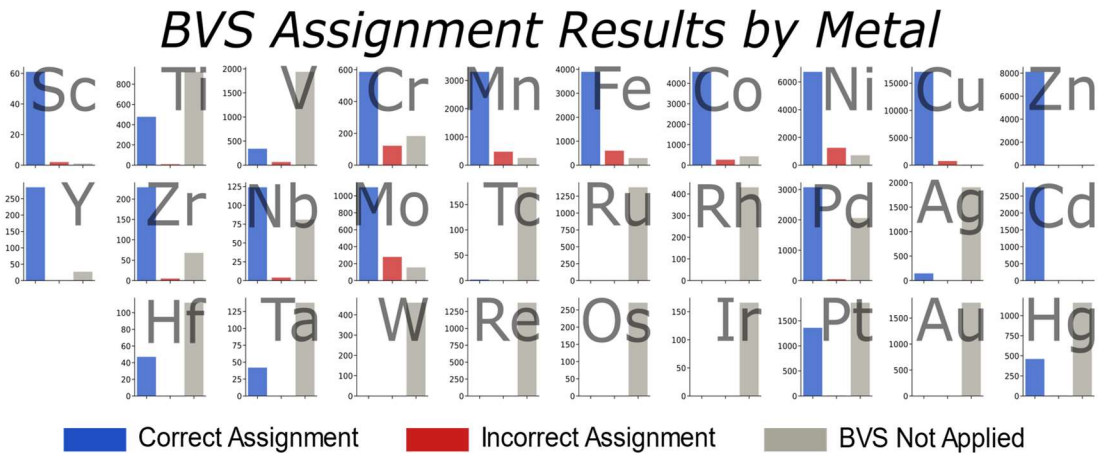


Figure 2.2: Application of the Bond Valence Sum Method by metal to test-set entries. Bars show relative success/failure and applicability of BVS for each of the 1st, 2nd and 3rd row transition metal atoms. BVS is not applicable when metal-ligand specific parameters are lacking for the common oxidation states of that metal.

Some measure of confidence in an oxidation state assignment can be obtained from (i) the agreement between different methods and (ii) whether any alerts are generated. The success in reproducing author-assigned oxidation states increases substantially over the data presented in Table 2.3 for the cases where an A or B confidence grade is obtained. Of the 36,080 entries with A assignment author-assigned oxidation states were reproduced in 99% of cases, with most of the "incorrect" assignments identifying structural or naming errors in the CSD (Figure 2.3).

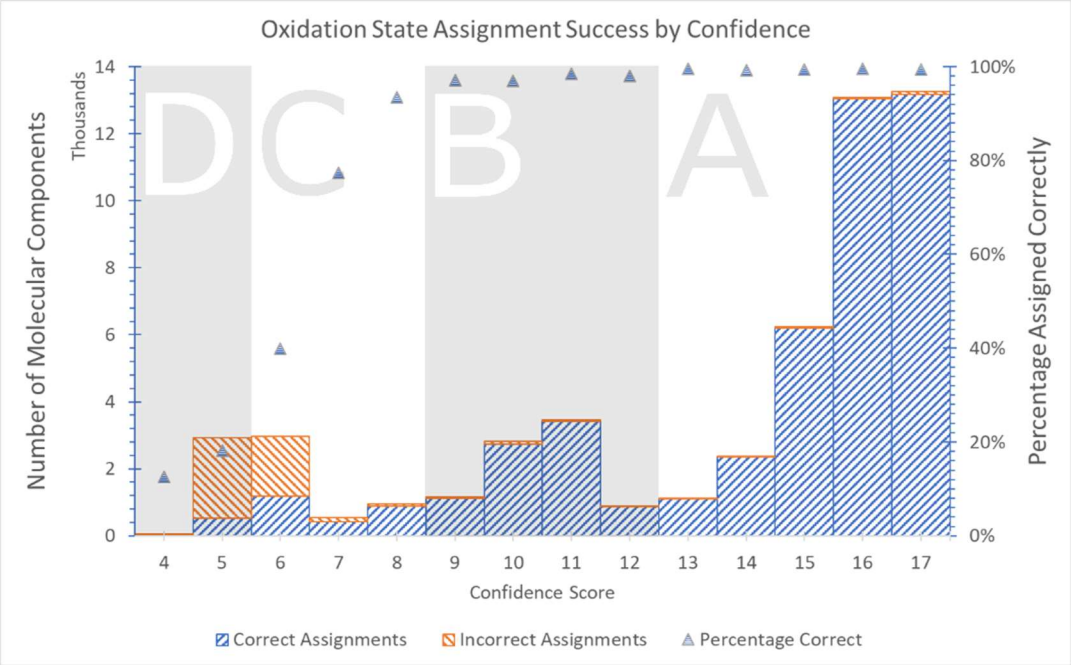


Figure 2.3: The success rate of oxidation state assignment grouped by confidence score bands. The raw data for this figure are available in the SI (Table S2.4).

2.3.2. Examples

Mononuclear complexes

In the Jahn-Teller distorted 6-coordinate Cu(2+) complex [diaqua-bis(pyrazine-2-carboxylato-N,O)-copper(II)], (refcode: BEYRAY03 [19], Figure 2.4.i) BVS parameters are available for Cu-O and Cu-N bonds for oxidation states +1, +2 and +3. As this covers the common copper oxidation states found in the look-up table, BVS is carried out and determines the oxidation state to be +2 with no warnings or errors. The aquo ligand is found to have a charge of zero, and the pyridine-carboxylate ligand a

charge of −1 by all three ligand charge methods. When the charges are compared to the fragment charge database no discrepancies are found. No errors or warnings are issued in the ligand charge-assignment procedure and the total ligand charge of −2 is consistent with the oxidation state assigned by the BVS method. The oxidation state of the Cu is thus assigned to +2 with a confidence score of 17 (A).

A similar process is observed for the nickel complex [bis(2-aminoacetato)-nickel(II) monohydrate] (refcode: LEPYOV [20], Figure 2.4.ii), where BVS can be applied to both the common (+2), as well as the less common (+3) oxidation states. A BVS valence of +2 is determined, and the total ligand charge calculations are consistent with this for all methods. This assignment is awarded the highest confidence score: 17 (A) with no discrepancies between these and previous ligand specific assignments.

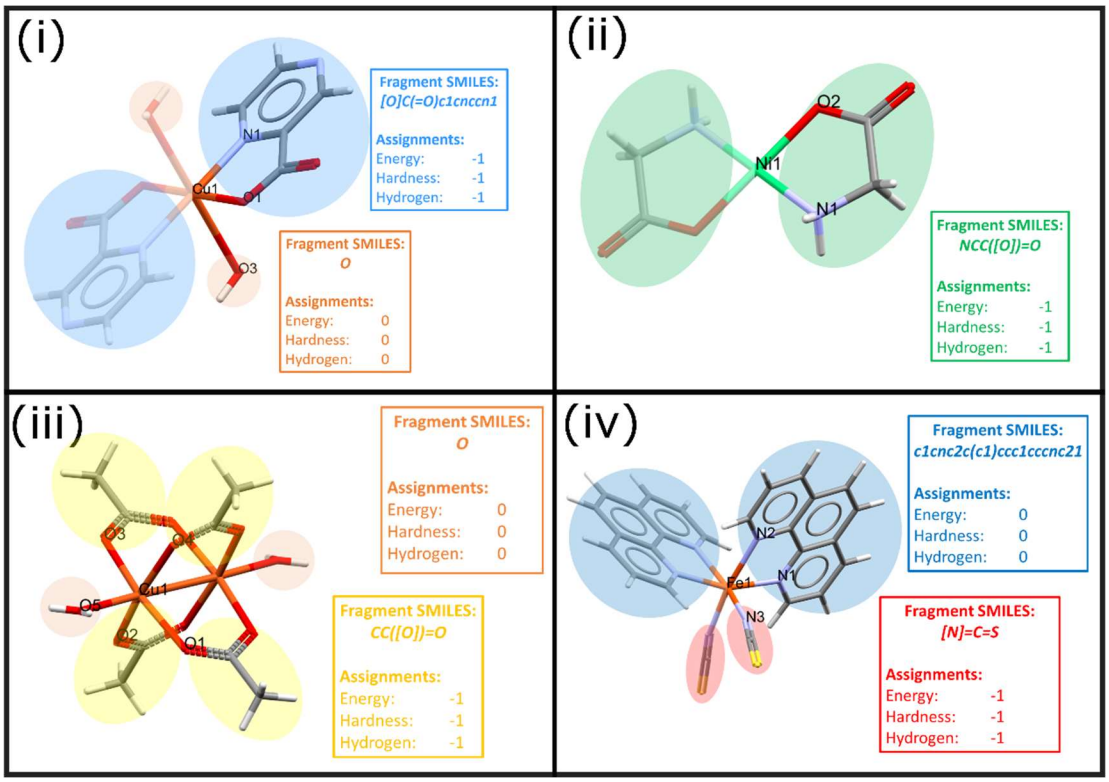


Figure 2.4: Oxidation state assignment examples. Boxes illustrate individual ligand fragments with corresponding SMILES notation and assignment method results demonstrated. Note only asymmetric unit metal-ligand atoms labelled for clarity. (i) Refcode: BEYRAY03; a copper(2+) structure with Jahn-Teller distortion. (ii) refcode: LEPYOV; a planar nickel(2+) complex. (iii) refcode: CUAQAC01; copper(2+) acetate with Cu-Cu bond as depicted in the CSD. (iv) refcode: KEKVIF; a family of Fe(2+) compounds that exhibit spin cross-over behaviour.

Metal-metal bonds

The identification of metal-ligand bonds has been based on inbuilt CSD functions for defining bonded atoms. The algorithm also generates bonds for short metal-metal distances. Metal-metal bonding is a widely-studied area of organometallic chemistry, but entries containing metal-carbon or metal-ligand multiple bonds were excluded from this study, and so many entries with metal-metal bonds would have been omitted on this basis. Nevertheless the CSD bonding criteria generate metal-metal bonds in coordination complexes such as the copper acetate dimer [tetrakis(μ^2-acetato)-diaqua-di-copper(+2),(refcode CUAQAC01,Figure 2.4.iii) in which the Cu…Cu distance is 2.619 Å. While metal-metal bond formation does not affect the ligand-charge procedures, the BVS method would fail because metal-metal bonds are not present in the bond valence parameter database used in this study.

Short metal-metal distances in coordination compounds are usually the result of the geometric demands of bridging ligands rather than genuine metal-metal bonding. M-M bonds have simply been omitted from the BVS calculations. The BVS calculation can then proceed as usual, yielding in the case of refcode CUAQAC01 a value of +2 for the copper oxidation state. The assignment is supported by each of the ligand charge methods. The confidence score is 17(A).

Spin cross-over complexes

The adoption of a high or low spin configuration affects metal-ligand bond distances and can influence oxidation state assignment via the BVS method. The crystal structures of many materials of interest in terms of spin cross-over (SCO) behaviour have been determined in multiple spin states, and occur in the CSD as refcode "families" where entries have the same six letter code but differ in the last two digits. While this test has focused on a single entry from each refcode family (see Section 3.1), to understand the role of SCO on oxidation state assignment, the process has been applied to a family of structures with both spin states present.

Complexes of Fe(+2) with nitrogen ligands have been widely investigated, and in this case both high spin and low spin bond valence parameters are available. For example, the refcode family for the iron complex [cis-bis(isothiocyanato)-bis(1,10-phenanthroline-N,N')-iron(+2)] (refcode: KEKVIF, Figure 2.4.iv), contains nine entries

with atomic co-ordinates. The ligand charge assignment methods produce the same result in each case, determining the thiocyanate and phen ligands to have charges of −1 and 0, yielding a metal oxidation state of +2. While the BVS method assigns an incorrect oxidation state of +3 for the low-spin entries the un-rounded values are all above 3.5 which generates a warning message. As a result of the discrepancies between BVS and ligand charge assignment methods, along with the warning in the BVS assignment, the low-spin complexes (such as Refcode: KEKVIF02) have a very low confidence score of 5(D). This situation occurs commonly with SCO families, and as such, SCO families are identifiable by large differences in confidence between entries.

For future assignments an extra parameter has been added to the confidence score which warns of potential refcode family issues. Where an entry is part of a refcode family, metal-bond distances are checked across the whole family when assigning oxidation states. If metal-ligand bond distances vary by more than 0.1 Å within the same refcode family, the confidence score is reduced by 4 and a warning issued.

Mixed valence polynuclear complexes

The assignment of oxidation states in polynuclear complexes is based on a combination of BVS and ligand charge methods. The Mn complex in Figure 2.5 (refcode: ZAVMEQ)[21] contains six unique metal centres with oxidation states between +2 and +4. The BVS method matched the literature values for all six metal centres, with no warnings or errors. The summation of ligand charges was −34 for all methods, which is consistent with the BVS total +34. In the ligand charge calculation for the pyridinyl-methanimine ligand there is a small (< 1 eV) difference in the hardness for charges of −1 and −3, which generates a warning. A warning is also produced for shallow energy curves for some fragments. The confidence score is 15(A).

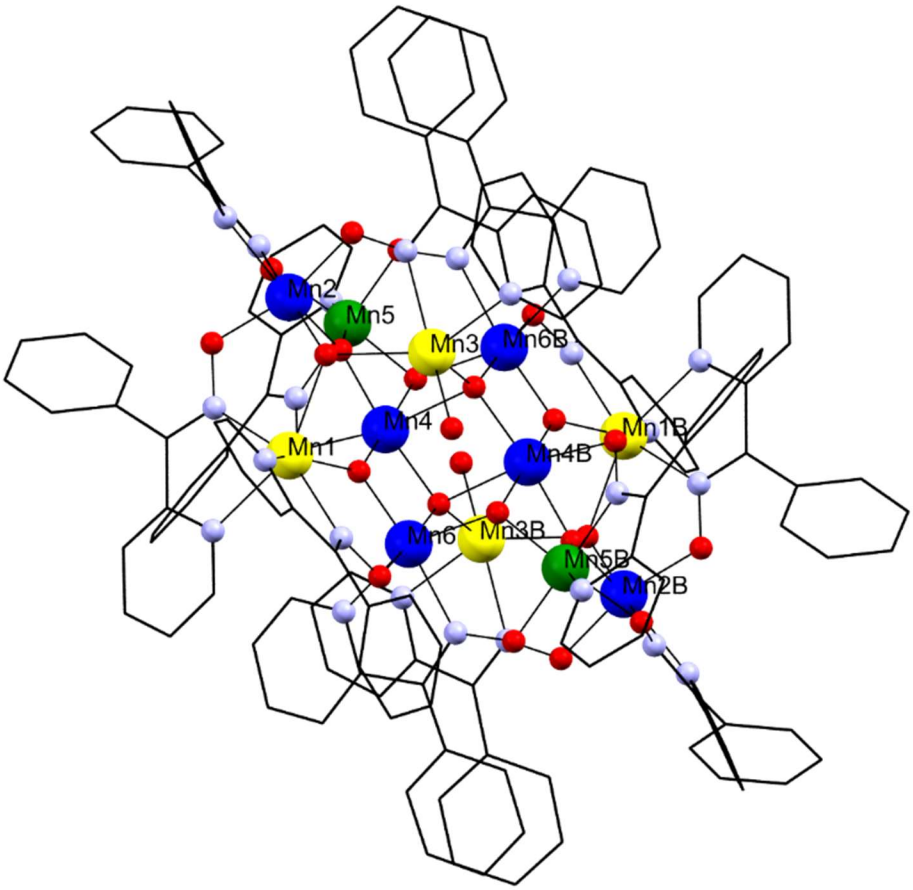


Figure 2.5: Mn12 structure (refcode: ZAVMEQ) containing six unique metal centres. Mn(2+) yellow, Mn(3+) blue, Mn(4+) green. H atoms omitted for clarity.

The CSD compound-naming conventions mean that a complex containing a single metallic element in multiple oxidation states would not have been part of the test data-set, and instead the procedures were validated manually by comparing assignments with those given in the corresponding publications. Approximately 50 complexes were examined over the course of this work and four errors were identified. In each case, the total of BVS assigned oxidation states did not match the ligand charge assignment methods and therefore all received a low confidence of 5/6.

Complexes with open-shell ligands

The closed-shell restriction applied during the ligand-charge calculations means that ligands with odd numbers of electrons are incorrectly treated. Complexes containing radical ligands were identified in the test data-set by comparing the author-assigned oxidation states with the allowed closed-shell charges for all ligands. A radical ligand is present if the named oxidation state does not match possible open shell charges. This was found to be the case for fewer than 2% of structures. Moreover, most of the instances involved a small set of common radicals. The SMILES formulae of identified radicals have been added to the ligand SQLite database. For example, nitroxide radicals are identified from SMILES string segments "CN([O])C", "cN([O])c", "cN([O])C" and "CN([O])c". It is additionally possible to add complete ligand specific SMILES manually for individual radical ligands where needed.

There is a tendency, in the case of open-shell ligands, for the energy and hydrogen placement charge-assignments to suggest values of −1, while hardness routinely suggests a charge of +1. BVS assignments are usually correct. This disagreement results in low confidence scores (typically D) where radical containing complexes are encountered for the first time. For example in the dinuclear 1,2,3,5-diselenadiazolyl nickel(+2) complex BARXID, BVS correctly assigns the oxidation state as +2 for both metal sites. The radical is given a charge of −1 according to energy and H-placement and +1 according to hardness. Overall the confidence score is 6 (C): 6 for the BVS method and 0s for all the ligand-charge criteria.

2.3.3. Demonstration of oxidation state specific data: the Jahn-Teller effect in Cu complexes

With atom specific valences now available, it is possible to limit some common geometric searches to specific oxidation states with relative ease. For example, the availability of atom-specific oxidation state data enables rapid collation of a list of Cu-ligand bond distances in Cu(+2) sites. Figure 2.6 shows a histogram of the distances obtained from copper sites with at least six short (>3.5 Å) Cu…N/O contacts where the oxidation state assignment has a confidence of A or B. The expected bimodal distribution between 1.8 and 2.8 Å shows elongation of M-L bond lengths for axial ligands. The plot enables an upper limit of about 2.8 Å to be placed on a Jahn-Teller axis in these complexes. A similar pattern is observed for Mn(3+) structures, with a bimodal distribution suggesting the same Jahn-Teller distortion out to 2.5 Å.

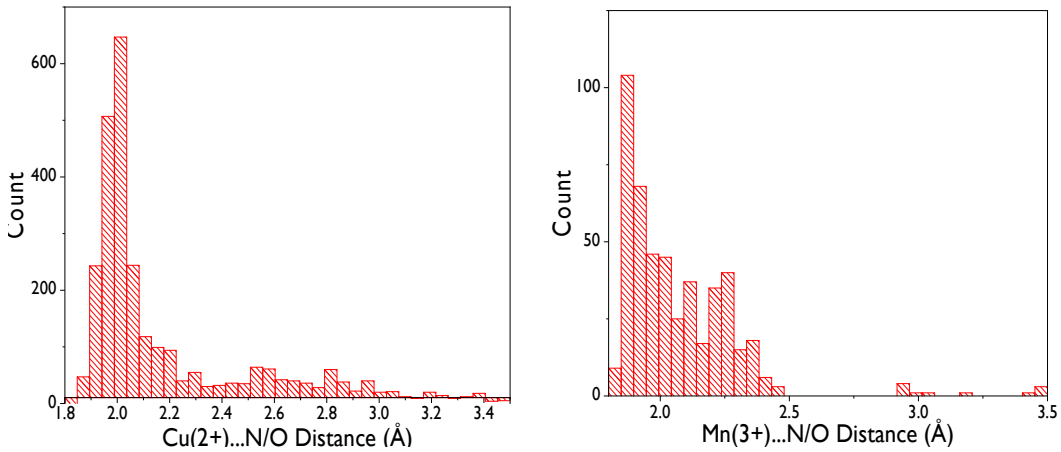


Figure 2.6: Distributions of Cu(2+)-N/O (left) and Mn(3+)-N/O (right) interatomic distances from the CSD.

2.4. Conclusions

The aim of the methods described in this chapter is to automate assignments of oxidation states to metal sites in mononuclear and polynuclear coordination complexes in the CSD. Each assignment is given a confidence score. Assignments with scores of A or B appear to be very reliable, yielding the correct assignment for 99% of cases during testing. Assignments with scores of C and D often represent special electronic or bonding situations, such as non-innocence associated with redox-active ligands, spin-state ambiguity or open-shell ligands. These cases still require manual checking. Experience over the course of this project suggests that the ultimate aim of completely automated oxidation state assignment without any manual intervention at all would be difficult or impossible to meet when based only on structural data.

The methods developed and investigated here will be implemented as part of the curation process of the CSD by expert scientific editors at the CCDC. In this manner, oxidation states where there is reliability in the assignment and/or clear pre-assignment by the author will be transferred straightforwardly into the curated CSD entry. The focussed attention from Scientific Editors can then be applied to the structures where the assignment is less reliable or indicates some very interesting or unusual chemistry.

A project is also currently underway to evolve the format of the CSD database and this will allow automated transfer of oxidation states from the compound in the entry (current) to be an atomic property on individual metal sites (future). The approaches described here will certainly help in that translation as well.

The availability of site-specific oxidation states as searchable criteria in the CSD should enable more targeted applications of the database in transition metal and materials chemistry. It should be possible, for example, to investigate how a metal and its oxidation state determines the deformability or structural flexibility of coordination; such information could be helpful in the design of metal templating reactions. Complexes with sites exhibiting unusual geometries might be susceptible to modification by high pressure or irradiation. The combination of oxidation-specific searching with motif-searching tools such as the Crystal Packing Feature in Mercury may find uses in research aiming to establishing the relationship between topology

and magnetic properties.[22,23] Finally, the SQLite ligand database could be extended to include a variety of properties such as conformational flexibility, pK_a, number of donor sites etc. that may be helpful in ligand design.

2.5. Available Stand-Alone Software

Although the methods described above are designed to work with curated entries in the Cambridge Database, a stand-alone script, named MRMOX, can be downloaded from the link http://www.crystal.chem.ed.ac.uk/software/mrmox. The script works through the MERCURY API to assign oxidation states with input from users' own cifs. A short set of installation and usage instructions is available in a read_me file in the download. The program will only work under WINDOWS with a fully functional version of MERCURY.

2.6. References

(1) Groom, C. R.; Bruno, I. J.; Lightfoot, M. P.; S. C. Ward. The Cambridge Structural Database. *Acta Cryst.* **2016,** *B72*, 171.

(2) Shields, G. P.; Raithby, P. R.; Allen, F. H.; Motherwell, W. D. S. The assignment and validation of metal oxidation states in the Cambridge Structural Database. *Acta Cryst.* **2000,** *B56* (3), 455.

(3) Davies, Daniel W.; Butler, K. T.; Isayev, O.; Walsh, A. Materials discovery by chemical analogy: role of oxidation states in structure prediction. *Faraday Discussions* **2018,** *211* (0), 553.

(4) Brown, I. D. *The Chemical Bond in Inorganic Chemistry: The Bond Valence Model*; Oxford University Press, 2016.

(5) Brown, I. D. IUCR, 2016; Vol. 2019.

(6) Housecroft, C. E.; Sharpe, A. G. *Inorganic Chemistry*; 3rd ed.; Pearson Prentice Hall: Harlow, Essex, UK, 2008.

(7) Cherif, I.; Abdelhak, J.; Zid, M. F.; Driss, A. Imidazolium trans-diaquadioxalatochromate(III) dihydrate. *Acta Cryst.* **2013,** *E69* (12), m667.

(8) Akbarzadeh Torbati, N.; Rezvani, A. R.; Safari, N.; Saravani, H.; Amani, V. Dichlorido(6,6'-dimethyl-2,2'-bipyridine-[kappa]2N,N')cobalt(II). *Acta Cryst.* **2010,** *E66* (10), m1284.

(9) Stewart, J. J. P., 2016.

(10) Stewart, J. J. P. Optimization of parameters for semiempirical methods VI: more modifications to the NDDO approximations and re-optimization of parameters. *Journal of Molecular Modeling* **2013,** *19* (1), 1.

(11) Pearson, R. G. The principle of maximum hardness. *Acc. Chem. Res.* **1993,** *26* (5), 250.

(12) Shields, G. P.; Raithby, P. R.; Allen, F. H.; Motherwell, W. D. S. The assignment and validation of metal oxidation states in the Cambridge Structural Database. *Acta Crystallographica Section B* **2000,** *56* (3), 455.

(13) Torzilli, M. A.; Colquhoun, S.; Doucet, D.; Beer, R. H. The interconversion of dichlorobis(N-n-propylsalicylaldimine)zinc(II) and bis(N-n-propylsalicylaldiminato)zinc(II). *Polyhedron* **2002,** *21* (7), 697.

(14) Rotondo, E.; Cusmano Priolo, F.; Bombieri, G.; Bruno, G. (N-{2-[2-(2-Ammonioethylamino)ethylamino]ethyl}salicylideneaminato-O,N,N',N'')perchloratocopper(II) perchlorate, [Cu(ClO4)(C13H22N4O)].ClO4. *Acta Cryst.* **1984,** *C40* (6), 960.

(15) Mezei, G.; Raptis, R. G. Effect of pyrazole-substitution on the structure and nuclearity of Cu(II)-pyrazolato complexes. *Inorg. Chim. Acta.* **2004,** *357* (11), 3279.

(16) Hipp, D. R.; 3.29.0 ed., 2019.

(17) Holler, S.; Tüchler, M.; Belaj, F.; Veiros, L. F.; Kirchner, K.; Mösch-Zanetti, N. C. Thiopyridazine-Based Copper Boratrane Complexes Demonstrating the Z-type Nature of the Ligand. *Inorg. Chem.* **2016,** *55* (10), 4980.

(18) Das, A.; Basuli, F.; Peng, S.-M.; Bhattacharya, S. Oxidation of Rhodium(I) by Hydroxamic Acids. Synthesis, Structure, and Electrochemical Properties of Bis(hydroxamate) Complexes of Rhodium(III). *Inorg. Chem.* **2002,** *41* (2), 440.

(19) Wang, G.-H.; He, R.-L.; Meng, F.-J.; Hu, N.-H.; Xu, J.-W. A polymorph of diaquabis(pyrazine-2-carboxylato-[kappa]2N1,O)copper(II). *Acta Cryst.* **2009,** *E65* (12), m1511.

(20) Wang, Z.-L. A monoclinic polymorph of bis(2-aminoacetato)nickel(II) monohydrate. *Acta Cryst.* **2006,** *E62* (10), m2546.

(21) Alexandropoulos, D. I.; Manos, M. J.; Papatriantafyllopoulou, C.; Mukherjee, S.; Tasiopoulos, A. J.; Perlepes, S. P.; Christou, G.; Stamatatos, T. C. "Squaring the clusters": a MnIII4NiII4 molecular square from nickel(ii)-induced structural transformation of a MnII/III/IV12 cage. *Dalton Trans.* **2012,** *41* (16), 4744.

(22) Macrae, C. F.; Bruno, I. J.; Chisholm, J. A.; Edgington, P. R.; McCabe, P.; Pidcock, E.; Rodriguez-Monge, L.; Taylor, R.; van de Streek, J.; Wood, P. A. Mercury CSD 2.0 - new features for the visualization and investigation of crystal structures. *J. Appl. Cryst.* **2008,** *41* (2), 466.

(23) Childs, S. L.; Wood, P. A.; Rodríguez-Hornedo, N.; Reddy, L. S.; Hardcastle, K. I. Analysis of 50 Crystal Structures Containing Carbamazepine Using the Materials Module of Mercury CSD. *Cryst Growth Des.* **2009,** *9* (4), 1869.

3. MrPIXEL: Execution of PIXEL Calculations via the MERCURY Interface

3.1. Introduction

3.1.1. Intermolecular Interactions in Crystal Structures

Intermolecular interactions control an enormous diversity of chemical and physical properties in solid materials including the phase adopted under a given set of applied conditions, the solubility and melting point, and thermodynamic properties such as lattice energy, hardness, thermal expansion, heat capacity and so on. The principal aim of many crystal structure determinations, particularly in fields such as crystal engineering and polymorphism research, is to understand relative phase stability and the significance of specific intermolecular contacts including 'structure directing' interactions such as the carboxylic acid dimer, nitro-iodo, and acid-pyridine contacts.[1] Use of these 'synthons' has been exploited in, for example, formation of co-crystals with active pharmaceutical ingredients with the aim of generating crystalline forms with improved performance.[2]

Intermolecular interactions are most commonly identified and ranked by assuming that short interatomic interactions characterise stabilising interactions, with the shortness of a contact defined relative to the sum of the van der Waals radii of the atoms in question. This method is not only quick and amenable to graphical visualisation, but use of a common set of radii such as Bondi's [3,4] 'prehistoric' [5] compilation or Alvarez's more recent and much more extensive set,[6] provides a unifying framework for discussion of intermolecular contacts.

Use of interatomic distances for interpreting crystal structures will tend to bias analyses towards those contacts in which interactions are mediated by specific atom-atom interactions. While hydrogen bonds, halogen bonds and the growing catalogue of related interactions are readily identifiable, stabilising contacts between non- or weakly- polar molecules, which are better understood in terms of whole molecule interactions, are harder to identify. The lack of a distinctive interatomic signature in van der Waals interactions (*i.e.* those dominated by dispersion) has led to their significance being unrecognised [7]. For example, the crystal structure of $MeNSOF_2$ contains no interatomic contacts at all when judged using van der Waals radii, yet still it is a solid with intermolecular energies similar to those found for medium-strength

hydrogen bonds and a lattice energy of a similar magnitude to acetic acid (−62 & −69 kJ mol^{-1}, respectively).[8,9] Focussing on prominent interactions can also give a misleading impression of the nature of an overall intermolecular contact. For example, out of 14 molecule-molecule contacts within the first coordination sphere of γ-glycine, six are destabilising, including two which are linked by hydrogen bonds.[10]

When considering thermodynamic stability, there are obvious advantages to working in Joules rather than Ångstroms, and there is a growing interest in interpreting crystal structures using whole molecule - whole molecule energies.[11-13] Quantum mechanical methods enable interaction energies to be computed to a very high level of accuracy, as illustrated by the *ab initio* calculation of the sublimation energy of benzene, can be very time consuming for large molecules.[14] While most quantum mechanical methods provide a single intermolecular energy, some, including symmetry adapted perturbation theory [15,16] break the energy down into constituent electrostatic, polarisation, dispersion and Pauli repulsion terms providing insight into the physical nature of an interaction. Though these methods are a gold standard in the field, they, too, are time consuming for large molecules when applied at the most accurate level.

3.1.2 CLP-Pixel

The Pixel method, which was originally devised by Gavezzotti,[17-19] adopts a semi-empirical approach in which molecules in a crystal structure are represented by blocks of electron and nuclear density sub-divided into small cubic volume elements referred to as *pixels*. The molecular electron densities are calculated *ab initio* using Gaussian [20] at the MP2 or B3LYP level, commonly with the 6-31G** basis set using a grid of spacing 0.08 Å. To save computer time the grid is 'condensed' into a coarser one, typically of dimension 0.32 Å. The electrostatic energy between two molecules can then be calculated by applying Coulomb's law to pairs of pixels from each molecule and then summing the values. A similar approach can be used for the polarisation, dispersion and repulsion terms, to achieve a total intermolecular energy broken down into physically meaningful contributions. Application of this approach to a cluster of molecules surrounding a central reference molecule enables the lattice energy to be evaluated. The most appropriate cluster radius depends on the size and nature of the molecules, but is typically between 12 and 20 Å. The accuracy of the

methods has been discussed by Chickos and Gavezzotti by comparing calculated sublimation energies with a large database of the experimental values.[21] Overall, the performance of the Pixel method is similar to periodic density functional theory in estimating sublimation enthalpies of organic solids, but at a fraction of the cost in terms of computing time.[22]

The Pixel method has been applied to numerous systems, such as in the quantitative investigation and description of synthons for crystal engineering.[7] Recent work using Pixel helped elucidate the role of intermolecular interactions and lattice energies for polymorphs of 5-methyl-2-[(2-nitrophenyl)-amino]-3-thiophenecarbonitrile (colloquially known as ROY) at high pressure.[23] Pixel has also been used to identify and rationalise the metastable form of glycolide,[24] explain the effect of chemical substitution on halogen bonding [25] and identify the features of racemic and homochiral polymorphs that make them thermodynamically competitive.[26]

The workflow of a Pixel calculation is shown in Figure 3.1: Process diagram for a Pixel calculation using the Pixel-C program within the CLP-Pixel package starting from the results of a crystal structure determination. Green boxes show files and blue boxes show processes. Intermolecular energies are sensitive to H-atom positions, and if the crystal structure was determined using X-rays the distances involving hydrogen atoms should be 'normalised' to the more accurate typical values seen by neutron diffraction. Certain other modifications may also be necessary (see Section 2.1. below). The atomic positions, their type (e.g. sp^2 or sp^3 C etc.) and the space group symmetry is defined in an initial set-up file (*.oeh*). A routine (Pixmt3) generates both the Pixel calculation file (*.inp*), and Gaussian job input file (*.gjf*) necessary to determine the electron density. Gaussian may be run locally or remotely to produce a cube-format electron density file (*.den*). Parameters for the calculation, e.g. those that control the pixel size and other parameters are stored in a separate *.par* file. The results of each calculation are stored in plain text *.pri* and *.mlc* files which report overall calculation results and individual dimer results, respectively.

The aim of this chapter is to describe a set of Python scripts that automate the process described above directly from the interface of the Mercury structure

visualisation software which is distributed with the Cambridge Structural Database (CSD). The scripts described make use of the CSD Python API. Pixel calculations are also accessible via the recently described MiCMoS package of computer programs, which brings together the AA-CLP, CLP-Pixel and CLPDyn procedures.[27] In addition, a procedure based on periodic electron densities is available in the program q-GRID.[28] Graphical visualisation of Pixel results can be accomplished with the processPIXEL procedure.[29]

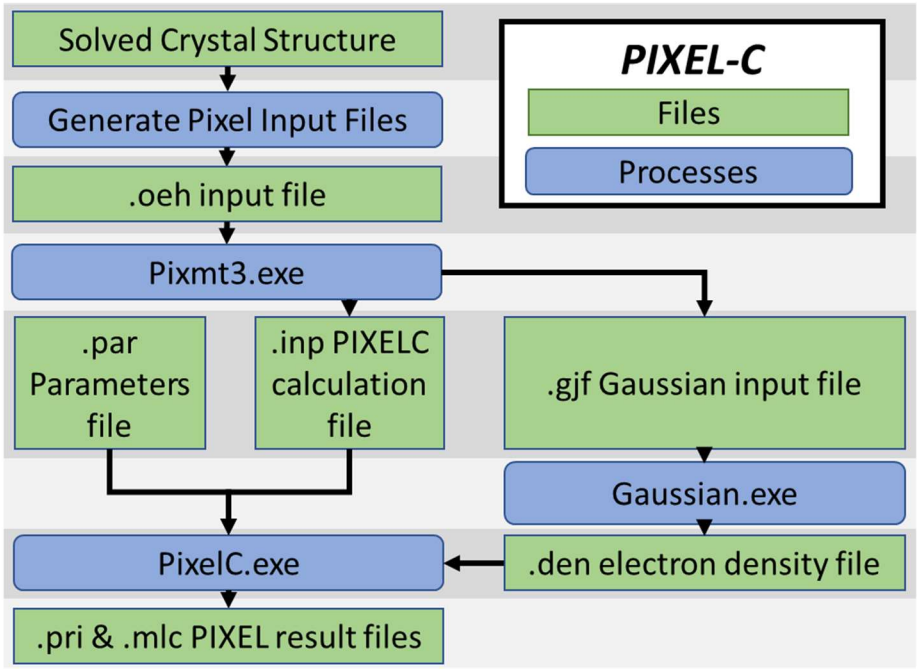


Figure 3.1: Process diagram for a Pixel calculation using the Pixel-C program within the CLP-Pixel package starting from the results of a crystal structure determination. Green boxes show files and blue boxes show processes.

The Program MrPIXEL consists of two elements. The first, SetupPixel, is executed from within Mercury, interprets the crystal structure and generates the input files required for the Pixel calculation. The second, MrPIXEL Console, handles all processes after initial file setup and displays the status of calculations through a graphical interface. A supplementary program, MrPIXEL Settings, is used to define default settings, file locations and so on.

3.2. Procedure

3.2.1. Modification of the Crystal Structure.

Before generating the necessary input files, structures may require some modification in order to satisfy the requirements of the Pixel program. The number of molecules in the asymmetric unit (*Z'*) is limited to two, and these must comprise complete molecules. Therefore, where a molecule occupies a special position, the space group symmetry of the crystal structure should be reduced to a description in which the asymmetric unit consists of whole molecules. The CSD's Mercury software can be used for this purpose, using the *"Change Spacegroup to Subgroup"* tool found in the *Edit* menu. For example, in the crystal structure of 2,2'-bipyridyl (CSD Refcode: BIPYRL), the molecule occupies the inversion centre in space group $P2_1/c$ with an asymmetric unit that consists of half the molecule. By reducing the space group to either *Pc* or $P2_1$ the inversion centre is removed, and the asymmetric unit contains a complete molecule. It should be noted however that in cases where Z' = 1.5, reduction of the symmetry results in Z' = 3, which cannot be accommodated in a Pixel lattice energy calculation. For example, in benzidine form III (CSD Recode: BENZIE02 also $P2_1/c$) one molecule occupies a general position and the other an inversion centre, and no choice of reduced symmetry will satisfy the conditions for Pixel [30]. In cases where symmetry lowering involves an origin shift it is advantageous to use a non-standard setting of the lower-symmetry space group to ensure that intermolecular relationships that do not involve the symmetry operations lost in the symmetry lowering are preserved exactly. There are, in short, a number of alternative strategies that may be used to lower symmetry, and for this reason no attempt has been made to implement a general procedure.

In addition, Pixel does not handle disordered structures. Where disorder is present, an ordered model should be constructed which contains just one component. Again, this is not handled automatically as users will usually wish to select which disorder component to keep. The required editing can be accomplished with the "*Edit Structure*" tool under the Mercury *edit* menu. Because the CSD Python API reads crystal structures upon loading into the window, the edited structure should be saved as a *.cif* and reloaded after any changes are made.

Once any necessary editing has been carried out, the SetupPixel script can be run from the Mercury CSD Python API menu. Upon selecting SetupPixel, the structure is read and interpreted, and the user is then prompted for some Pixel calculation parameters. This includes a calculation name (defaulted to the *.cif* title), whether to normalise hydrogen positions (required for conventionally-modelled structures derived from X-ray data), and the charge and spin-multiplicity of each component.

3.2.2. Generating the .oeh input file.

The user is given the option to generate the initial Pixel files only or pass the task onto MrPIXEL Console so that the entire calculation can be run automatically. In either case, SetupPixel first generates the Pixel input *.oeh* file. Information relating to crystal symmetry, cell parameters, and atomic positions are taken from the structure as interpreted by Mercury. Atomic type indicators, as defined by Pixel (see SI Table S3.1), use similar definitions as the Tripos *.mol2* SYBYL typing. These atomic type indicators can therefore be determined by a translation of SYBYL to Pixel types using a look-up table. This is particularly important for carbon atoms, where Pixel assigns atomic polarisabilities based on the bonding environment. SYBYL assignments are based on the results of the CSD *auto_edit* structure tool, which adds atom and bond-type descriptors to the structure. To store this information, and to allow users to inspect the assignments, the resulting structure is saved as a *.mol2* file.

It is also possible at this stage to select the level of theory and basis set for the Gaussian calculations. For organic molecules containing atoms with atomic numbers up to bromine, MP2/6-31G** is usually used; B3LYP/6-31G** is used for first row transition metal complexes. Different levels of theory as well as additional Gaussian set-up settings, are stored as a history that allows the user to select the correct job line as required.

3.2.3. Running Pixmt3

Once written, the *.oeh* file is passed to the CLP-Pixel Pixmt3 routine which generates input files for Gaussian and Pixel. The electron density step size and van der Waals radius parameter (rvdw) values may be in specified here or left as a default value set in the settings. The electron density step size denotes the dimensions of each pixel cube calculated in Gaussian. The default values work well for elements up to bromine,

but the step size should be reduced to 0.06 Å or lower for structures involving heavier atoms [25]. To reduce computational time, Pixel calculations are run using 'super-pixels' of n^3 pixels, where n is known as the 'condensation level'; n usually set to 4 so that a step size of 0.08 Å produces super-pixel cubes of dimension 0.32 Å. For most purposes this condensation level is adequate, but it can be changed in the MrPIXEL Settings. Any necessary changes to the file output by Pixmt3, e.g. to the basis set, are made by MrPIXEL Console at this stage.

3.2.4. Generating Gaussian electron density files

MrPIXEL Console accommodates Gaussian calculations performed both locally and on remote cluster installations. For remote jobs, MrPIXEL Console interfaces with clusters through the Python SSH module Paramiko. Paramiko enables the scripts to connect securely to remote clusters via an SSH key combination. The username and password are needed only for initial set-up of the key files on both the local and remote locations. Future connections match these key files. The details of submission and retrieval depend on local cluster type and administration policy. The scripts used in Edinburgh are included in the MrPIXEL package but it is expected that these will usually need some modification in other locations.

The cluster address and folder locations can be specified in the MrPIXEL Settings. In the system implemented in Edinburgh Gaussian jobs are submitted using a Bash script, and a template Bash file should be included in the *PIXEL\Batch* folder on the local system. MrPIXEL Console produces an edited copy of this for each calculation, replacing the entry *"Name.gjf"* to the job file name. Calculations are checked periodically when running MrPIXEL Console and the density files downloaded when Gaussian jobs are complete. A job is deemed complete when the required electron density 'cube' file is retrieved.

3.2.5. Running the Pixel calculation

Following retrieval of the density file, the Pixel calculation is called by MrPIXEL. Pixel calculations are carried out using the Pixel-C module of the CLP-Pixel suite, using all available CPU resources on a single core. It is not recommended to run calculations concurrently that equal or exceed the core count on a user's machine. The user may therefore specify in the settings a maximum number of cores available to Pixel-C

tasks in the settings menu, which will result in MrPIXEL queueing tasks that exceed this limit until a free core is available. MrPIXEL Console reports the completion of Pixel tasks in the GUI and provides functionality to view the interactions as sorted tables in a text viewer under the MrPIXEL Console -> Structure menu.

3.3. Examples

3.3.1. The First Coordination Sphere of γ-Glycine

A straightforward calculation may be demonstrated using the structure of γ-glycine determined by neutron diffraction (CSD Refcode GLYCIN16) [31]. The structure contains one molecule in the asymmetric unit which occupies a general position in the space group $P3_2$. No modifications to the H-atom positions or space group are necessary, and SetupPixel can be called without any manual modification (Figure 3.2).

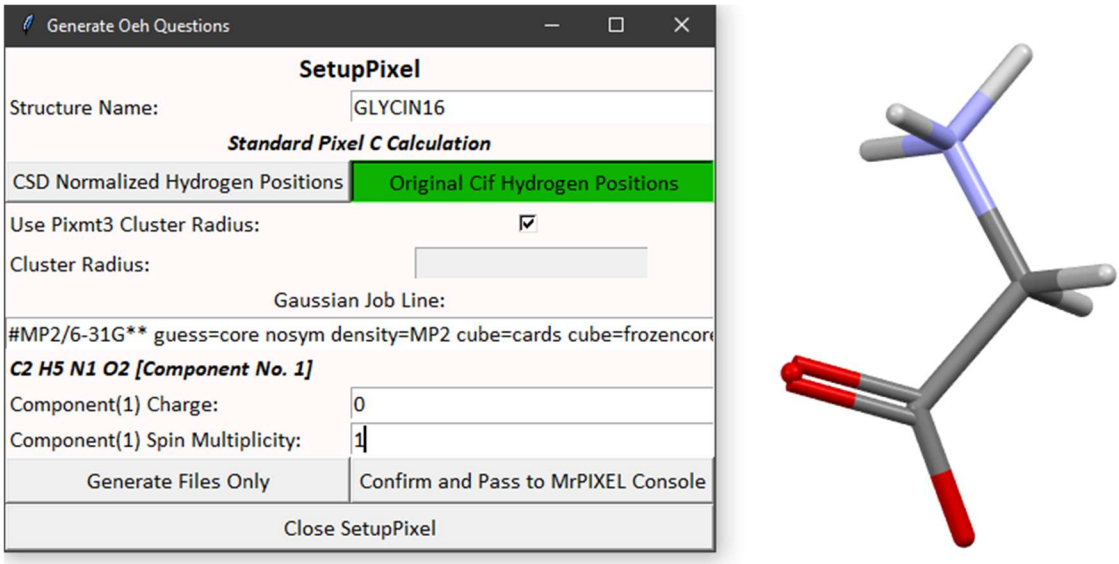


Figure 3.2: SetupPixel window for the GLYCIN16 CSD structure showing typical settings.

The SetupPixel menu defines settings for the Pixel calculation. The first line describes the job title and may be changed as required; the default is taken from the structure data title in the CIF. For CSD entries this comes from the CSD refcode. The next line defines the job type. If $Z' \leq 2$ this will be a standard Pixel calculation. Where more than two molecules occupy the asymmetric unit, a series of separate Pixel calculations are carried out with pairs of components (see below). The user then selects whether to normalise hydrogen positions (as is typical for X-ray data). In this example, GLYCIN16 was determined using neutron data and so normalisation is not

needed. The next option determines whether the cluster radius should be as determined by Pixmt3 or defined by the user. For the present calculation, our interest is only in the dimer energies of the first co-ordination sphere, and the Pixmt3 cut-off (14 Å) is more than adequate. Finally, the charges and spin multiplicity values are required for each molecule. In this case the defaults (charge = 0, spin multiplicity = 1) are correct for the glycine molecule.

The calculation is passed to MrPIXEL. The user is prompted for the electron density grid size for the cube-format file generated in the Gaussian calculation. A Gaussian job is then submitted, and its progress monitored; when complete the cube file is downloaded and the Pixel calculation initiated.

Pixel returns a lattice energy of −235.9 kJ mol^{-1}, though it should be noted that glycine is zwitterionic in the solid state and contact energies at the cut-off radius of 14 Å still have interaction energies in excess of 2 kJ mol^{-1} as a result of long range electrostatic interactions. The experimental lattice energy of glycine is between −136 and −139 kJ mol^{-1}, the large difference with the calculated value reflecting the transfer of a proton between the ammonium and carboxylate groups which occurs in the gas phase.[9] The lattice energy of glycine neglecting the proton transfer has been estimated to be −290(8) kJ mol^{-1}.[32]

Molecules in the first coordination sphere can be identified from non-zero values of the repulsion or dispersion energies, which are very short-range interaction terms. There are 14 molecules in the first coordination sphere (Table S3.2), the strongest interaction being the head-to-tail H-bond formed between the ammonium and carboxylate groups which forms a *C*(5) chain along the *c*-axis. The interaction is dominated by the electrostatic term (−119.2 kJ mol^{-1}), with a smaller contribution from dispersion, as is typical for H-bonds. The pattern of interactions can also be visualised with the ProcessPixel software.[29]

3.3.2. The Lattice Energy of Ethylene

The space group of the crystal structure of ethylene (ETHLEN10) is $P2_1/n$ with the molecule located on an inversion centre ($Z' = ½$) [33]. The space group symmetry needs to be reduced to a $Z' = 1$ description, either in $P2_1$ or *P*c. This step should be carried out before setting up the Pixel calculation and can be accomplished in Mercury

as described above (Section 2.1). The updated structure should be saved as a CIF which can be opened within the same Mercury window before running SetupPixel.

The structure was determined using X-rays, and so H-atoms should be normalised. The influence of the cut off radius on the lattice energy and number of interactions is shown for ethylene in Figure 3.3. The lattice energy hardly changes beyond 10 Å because the influence of electrostatic contributions is very low. In practice validation of the cut off can be carried out after Pixel calculation is complete by checking that interaction energies at the longest distances are zero.

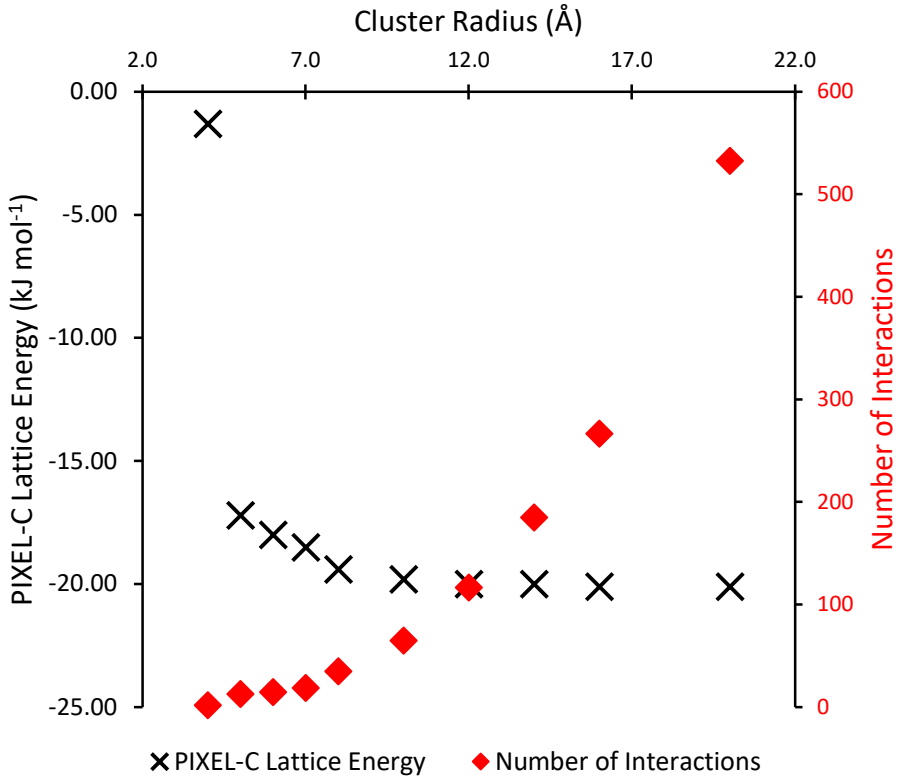


Figure 3.3: Influence of cluster radius on lattice energy for ETHLEN10 as calculated by Pixel (black). The red points show the number of interactions considered at each choice of cluster radius.

The calculated lattice energy is −22.4 kJ mol^{-1}, which compares to experimental values of between −20.2 and −25.2 kJ mol^{-1}. The breakdown of the contributions to the lattice energy are (in kJ mol^{-1}) −5.8 for the electrostatic energy, −1.5 for polarisation, −27.6 for dispersion and 12.5 for repulsion. The first coordination contains 12 contacts, the interactions having a fairly uniform distribution of energies.

3.3.3. A Transition Metal Complex

Parameterisation and application of the Pixel method to transition metal complexes has been carried out using electron densities calculated using B3LYP/6-31G** calculations,[34,35] and this is specified when running SetupPixel as shown in Figure S3.1.iii in the SI.

For the low-spin Mn(+3) containing salt $[Mn(cyclam)(CN)_2]ClO_4$ (AFAROO) the spin multiplicity is 3 for the cation, while the charges on the cation and anion are +1 and −1.[36] The energies from Pixel are classified according to whether they are cation-cation, cation-anion, anion-cation and anion-anion interactions. The composition of the first coordination sphere of the cation is shown in Table S3.4 and in Figure 3.4.i, where the central cation, labelled M, makes contacts to ten other cations and six perchlorate anions, labelled A1, A2 etc. Cations M3 to M8 are distributed in a distorted cube about the central cation. The interactions are dominated by dispersion with total energies in the range −9.1 to −19.1 kJ mol^{-1}. Two pairs of anions lie at the edges of cube, with the remaining two anions (A3 and A4) occupying the opposite faces. Topologically it is similar to the CoO structure.[37] The strongest contacts (−66.5 kJ mol^{-1}) are formed to two cations (M1 and M2) located in the top and bottom faces of the cube, connected by pairs NH…NC H-bonds between the cyclam and cyano ligands (Figure 3.4.ii). Unusually for an ionic material, electrostatics appear to play a minor role, while the most strongly stabilising electrostatic interaction (to M1 and M2) is formed between two cations. These results reflect the distribution of a single positive charge over a relatively large cation and the retention of a significant negative electrostatic potential in the region of the cyano ligand, which Mulliken analysis shows to carry a charge of approximately −0.5 e.

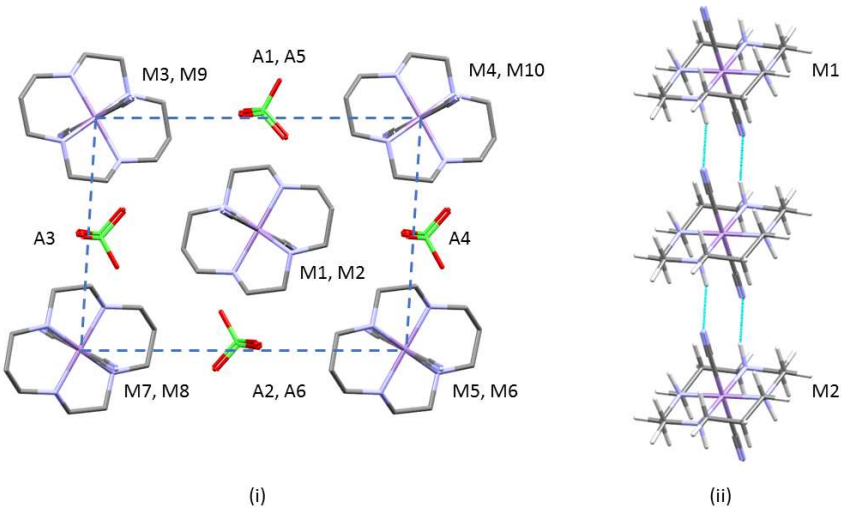


Figure 3.4. (i) The first coordination sphere of the cations in the Mn complex AFAROO. The dashed line shows the top face of the distorted cube referred to in the text. (ii) H-bonding forming chains of cations.

3.3.4. Pixel Calculations when *Z*' > 2

For structures with more than two molecules in the asymmetric unit, a standard Pixel calculation is not possible (see Section 2.1). It is, however, possible to run multiple Pixel calculations to obtain individual dimer energies in the structures by consideration of substructures consisting of all possible pairs of molecules in the asymmetric unit. SetupPixel will recognise such structures, as shown in S.I. Figure S3.1iv., notifying the user, and run iterations of Pixel to generate all the dimer energies out to the set cut-off range. The setup of calculations is the same as usual, but the output folder will contain calculation files for each possible combination of molecules. Note that the lattice energies obtained in these calculations are meaningless, and so only relatively short cut off radii are required.

This process can be applied to the structure of acetoxime (CSD refcode ACEOXM01, $Me_2C{=}NOH$),[38] which has three molecules in the asymmetric unit in space group $P\bar{1}$. The output for this example contains three Pixel-C calculation results (corresponding to interactions between molecules labelled a and b, a and c and b and c). The structure, which at 220 K has unit cell dimensions a = 7.01, b = 10.48, c = 10.58 Å, α = 60.5°, β = 79.6°, γ = 83.5° appears to be a distorted version of a hexagonal room temperature phase which forms in $P6_3/m$ with dimensions a = 10.61, c = 7.02 Å [39]. The first coordination sphere of each of the three molecules in the asymmetric unit contains 12 molecules, consisting of a central layer in which each

molecule is surrounded by six others generated by lattice translations. Layers above and below are related to the central molecule by inversion operations and to each other by lattice translations along *a* (Figure 3.5). Overall the arrangements have the characteristic ABAB… layer stacking of hexagonal close packing. The HCP topology allows equivalent contacts to be identified and compared (Table S3.5). The HCP arrangement is distorted in the parent phase by the non-spherical geometry of the molecules and H-bonding between the members of the asymmetric unit, but each of the contacts in horizontal rows in Table S3.5 would have been equivalent in the parent phase and show still further variation. The H-bonds in the first two rows of the table are dominated by the electrostatic contribution and show less variation than the interactions between the layers (bottom six rows) which are dominated by dispersion, illustrating the flexible character of dispersion interactions. The sums of the contacts in the three E_{TOT} columns are −124.8, −125.3 and −125.5 kJ mol^{-1}, demonstrating the mutual compensation of the distortions that occur about each molecule.

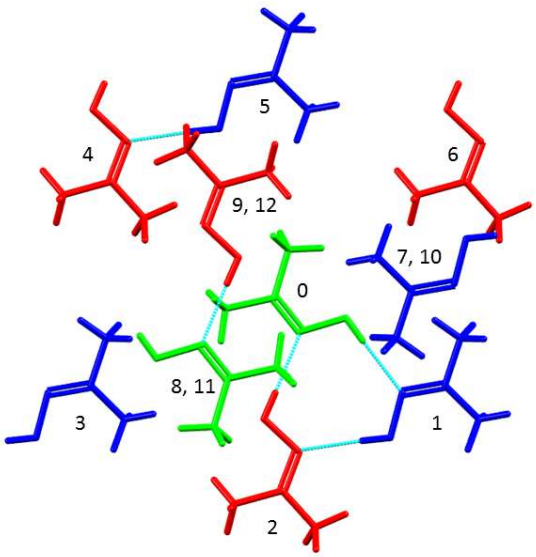


Figure 3.5: The first coordination sphere of molecule 'a' in acetoxime shown in green with molecules b and c shown in blue and red. The central reference molecule is labelled '0', with other contacts being labelled in the same order as Table S3.5. Molecule 7 and 10 etc. superimpose in this projection along **a**.

3.4. Conclusions and Program Availability

The availability of accurate, semi-empirical methods such as PIXEL and CrystalExplorer [13] for the calculation of intermolecular interaction energies in crystal structures provides thermodynamic insight into the intermolecular interactions which drive and determine crystal structure formation. They can be used to help interpret individual crystal structures, to compare the structures of different polymorphs, cocrystals and solvates and to quantify the effect of chemical substitution on interactions in a series of related materials. They are broadly applicable to a range of different compounds, rather than being limited to certain classes as are some molecular mechanics methods. Compared to fully *ab initio* quantum mechanical methods, they are not only very extremely fast, but they also provide a breakdown of intermolecular energies into chemically meaningful terms.

Long-standing methods for understanding crystal structures, such as the use of van der Waals radii to identify stabilising contacts, provide a way to identify atom-atom contacts and instantly place their distances in the context of similar interactions.[40] While the speed of such calculations is likely to ensure they will remain the first step of most crystal structure analyses, the calculation of energies is highly complementary and enriches the information content of a crystal structure. By emphasising molecule-molecule over atom-atom interactions it also simplifies the analysis of crystal structures by simply reducing the volume of numerical data that need to be considered.

The aim of MrPIXEL is to facilitate the development of a purely structural view of intermolecular interactions and crystal packing into a more fundamental thermodynamic one. Once the program is installed and set up the entire process of a PIXEL calculation can be carried out with the minimum of effort from the interface of Mercury. The code is open source and freely available from http://www.crystal.chem.ed.ac.uk/software/mrpixel. The package includes the programs Pixel-C and Pixmt3 from Pixel-CLP, the full version of which can now be downloaded from http://www.crystal.chem.ed.ac.uk.

3.5. References

(1) Mukherjee, A. Building upon Supramolecular Synthons: Some Aspects of Crystal Engineering. *Cryst. Growth Des.* **2015,** *15* (6), 3076.

(2) Aakeröy, C. B.; Chopade, P. D.; Desper, J. Avoiding "Synthon Crossover" in Crystal Engineering with Halogen Bonds and Hydrogen Bonds. *Cryst. Growth Des.* **2011,** *11* (12), 5333.

(3) Bondi, A. van der Waals Volumes and Radii. *J. Phys. Chem.* **1964,** *68* (3), 441.

(4) Bondi, A. van der Waals Volumes and Radii of Metals in Covalent Compounds. *J. Phys. Chem.* **1966,** *70* (9), 3006.

(5) Gavezzotti, A. The "sceptical chymist": intermolecular doubts and paradoxes. *CrystEngComm* **2013,** *15* (20), 4027.

(6) Alvarez, S. A cartography of the van der Waals territories. *Dalton Trans* **2013,** *42* (24), 8617.

(7) Dunitz, J. D.; Gavezzotti, A. Supramolecular Synthons: Validation and Ranking of Intermolecular Interaction Energies. *Cryst. Growth Des.* **2012,** *12* (12), 5873.

(8) Mews, R.; Parsons, S. Comparison of intermolecular interactions in two phases of MeNSOF2. *Zeitschrift für Kristallographie - Crystalline Materials* **2014,** *229*.

(9) Chickos, J. S.; Acree, W. E. Enthalpies of Sublimation of Organic and Organometallic Compounds. 1910–2001. *J. Phys. Chem. Ref. Data* **2002,** *31* (2), 537.

(10) Moggach, S. A.; Marshall, W. G.; Rogers, D. M.; Parsons, S. How focussing on hydrogen bonding interactions in amino acids can miss the bigger picture: a high-pressure neutron powder diffraction study of ε-glycine. *CrystEngComm* **2015,** *17* (28), 5315.

(11) Dunitz, J. D.; Gavezzotti, A. Molecular Recognition in Organic Crystals: Directed Intermolecular Bonds or Nonlocalized Bonding? *Angewandte Chemie International Edition* **2005,** *44* (12), 1766.

(12) Dunitz, J. Intermolecular atom-atom bonds in crystals?This essay is commented on by Thakur, Dubey & Desiraju [IUCrJ (2015), 2, 159-160] and Lecomte, Espinosa & Matta [IUCrJ (2015), 2, 161-163]. *IUCrJ* **2015,** *2* (2), 157.

(13) Mackenzie, C. F.; Spackman, P. R.; Jayatilaka, D.; Spackman, M. A. CrystalExplorer model energies and energy frameworks: extension to metal coordination compounds, organic salts, solvates and open-shell systems. *IUCrJ* **2017,** *4* (5), 575.

(14) Yang, J.; Hu, W.; Usvyat, D.; Matthews, D.; Schütz, M.; Chan, G. K.-L. Ab initio determination of the crystalline benzene lattice energy to sub-kilojoule/mole accuracy. *Science* **2014,** *345* (6197), 640.

(15) Hohenstein, E. G.; Sherrill, C. D. Wavefunction methods for noncovalent interactions. *WIREs Computational Molecular Science* **2012,** *2* (2), 304.

(16) Szalewicz, K. Symmetry-adapted perturbation theory of intermolecular forces. *WIREs Computational Molecular Science* **2012,** *2* (2), 254.

(17) Gavezzotti, A. Calculation of lattice energies of organic crystals: the PIXEL integration method in comparison with more traditional methods. *Zeitschrift für Kristallographie - Crystalline Materials* **2005,** *220* (5/6), 499.

(18) Gavezzotti, A. Efficient computer modeling of organic materials. The atom–atom, Coulomb–London–Pauli (AA-CLP) model for intermolecular electrostatic-polarization, dispersion and repulsion energies. *New J. Chem.* **2011,** *35* (7), 1360.

(19) Gavezzotti, A. *Molecular Aggregation*; Oxford University Press: Oxford, U.K., 2007.

(20) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.et al. Wallingford, CT, 2016.

(21) Chickos, J. S.; Gavezzotti, A. Sublimation Enthalpies of Organic Compounds: A Very Large Database with a Match to Crystal Structure Determinations and a Comparison with Lattice Energies. *Cryst. Growth Des.* **2019,** *19* (11), 6566.

(22) Maschio, L.; Civalleri, B.; Ugliengo, P.; Gavezzotti, A. Intermolecular Interaction Energies in Molecular Crystals: Comparison and Agreement of Localized Møller–Plesset 2, Dispersion-Corrected Density Functional, and Classical Empirical Two-Body Calculations. *J. Phys. Chem. A* **2011,** *115* (41), 11179.

(23) Funnell, N. P.; Bull, C. L.; Ridley, C. J.; Capelli, S. Structural behaviour of OP-ROY at extreme conditions. *CrystEngComm* **2019,** *21* (30), 4473.

(24) Hutchison, I. B.; Delori, A.; Wang, X.; Kamenev, K. V.; Urquhart, A. J.; Oswald, I. D. H. Polymorphism of a polymer precursor: metastable glycolide polymorph recovered via large scale high-pressure experiments. *CrystEngComm* **2015,** *17* (8), 1778.

(25) Carlucci, L.; Gavezzotti, A. A quantitative measure of halogen bond activation in cocrystallization. *PCCP* **2017,** *19* (28), 18383.

(26) Dunitz, J. D.; Gavezzotti, A. Proteogenic Amino Acids: Chiral and Racemic Crystal Packings and Stabilities. *J. Phys. Chem. B* **2012,** *116* (23), 6740.

(27) Gavezzotti, A.; Lo Presti, L.; Rizzato, S. Mining the Cambridge Database for theoretical chemistry. Mi-LJC: a new set of Lennard-Jones–Coulomb atom–atom potentials for the computer simulation of organic condensed matter. *CrystEngComm* **2020**, DOI:10.1039/D0CE00334D 10.1039/D0CE00334D.

(28) de Klerk, N. J. J.; van den Ende, J. A.; Bylsma, R.; Grancic, P.; de Wijs, G. A.; Cuppen, H. M.; Meekes, H. q-GRID: A New Method To Calculate Lattice and Interaction Energies for Molecular Crystals from Electron Densities. *Cryst. Growth Des.* **2016,** *16* (2), 662.

(29) Bond, A. processPIXEL: a program to generate energy-vector models from Gavezzotti's PIXEL calculations. *J. Appl. Cryst.* **2014,** *47* (5), 1777.

(30) Rafilovich, M.; Bernstein, J. Serendipity and Four Polymorphic Structures of Benzidine, C12H12N2. *J. Am. Chem. Soc.* **2006,** *128* (37), 12185.

(31) Kvick, A.; Canning, W. M.; Koetzle, T. F.; Williams, G. J. B. An experimental study of the influence of temperature on a hydrogen-bonded system: the crystal structure of [gamma]-glycine at 83 K and 298 K by neutron diffraction. *Acta Cryst.* **1980,** *B36* (1), 115.

(32) Raabe, G. The Use of Quantum Chemical Semiempirical Methods to Calculate the Lattice Energies of Organic Molecular Crystals. Part I: The Three Polymorphs of Glycine *Z. Naturforsch.* **1999,** *A54*, 611.

(33) Nes, G. J. H.; Vos, A. Single-crystal structures and electron density distributions of ethane, ethylene and acetylene. III. Single-crystal X-ray structure determination of ethylene at 85 K. *Acta Cryst.* **1979,** *B35* (11), 2593.

(34) Maloney, A. G. P.; Wood, P. A.; Parsons, S. Intermolecular interaction energies in transition metal coordination compounds. *CrystEngComm* **2015,** *17* (48), 9300.

(35) Maloney, A. G. P.; Wood, P. A.; Parsons, S. Use of the PIXEL method to investigate gas adsorption in metal–organic frameworks. *CrystEngComm* **2016,** *18* (18), 3273.

(36) Mossin, S.; Sorensen, H. O.; Weihe, H. trans-Bis(cyano-[kappa]C)(1,4,8,11-tetraazacyclotetradecane-[kappa]4N)manganese(III) perchlorate, a low-spin manganese(III) complex. *Acta Cryst.* **2002,** *C58* (4), m204.

(37) Tombs, N. C.; Rooksby, H. P. Structure of Monoxides of some Transition Elements at Low Temperatures. *Nature* **1950,** *165* (4194), 442.

(38) Parsons, S.; Pu, W.; Ramage, R.; Wood, P. A. *CSD Communication* **2004**.

(39) Bierlein, T. K.; Lingafelter, E. C. The crystal structure of acetoxime. *Acta Cryst.* **1951,** *4* (5), 450.

(40) Thakur, T. S.; Dubey, R.; Desiraju, G. R. Intermolecular atom-atom bonds in crystals - a chemical perspectiveThis essay and that by Lecomte et al. (2015) comment on Dunitz (2015). *IUCrJ* **2015,** *2* (2), 159.

4. Revealing Cooperativity Pathways in Spin Crossover Complexes

4.1. Introduction

In octahedral complexes in which the metal has a d^4-d^7 electronic configuration, the occupation of e_g and t_{2g} orbitals is directed by the nature of the ligand field. Weak field ligands result in a small energy difference between t_{2g} and e_g orbitals (Δ_O), while strong field ligands lead to a larger Δ_O. This leads to the possibility of two electron distributions dependant on the size of Δ_O relative to the spin pairing energy (E_P). Where E_p is greater than Δ_O, d-electrons are distributed between the t_{2g} and e_g orbitals before pairing electrons to form a high spin configuration (HS). In contrast, where E_p is smaller than Δ_O, electrons pair in the t_{2g} orbitals before occupying the higher energy e_g orbitals in a low spin (LS) state. [1,2]

In spin crossover (SCO) complexes, the values of Δ_O and E_p energies are similar, allowing a complex to exist as either HS or LS dependent on the amount of energy applied to a system in the form of temperature, pressure, or light. Such complexes can be reversibly switched between spin states, resulting in different magnetic, optical or structural properties. In thermally promoted spin crossover the low-spin state is enthalpically favoured at low temperature, whereas the high spin state is entropically favoured at high temperature. SCO do not typically occur at a sharp, well-defined temperature, and instead occur over a range of temperature. Moreover, even when a transition is first order, its sharpness depends on the compound and its phase. The transition temperature ($T_{½}$), where the occupancies of molecules in the HS and LS states are equal, can be measured by several techniques including magnetic susceptibility, Mössbauer and Raman spectroscopies.[3-5] The abruptness of the transition, also known as the transition width, represents the temperature range over which the crossover occurs and has been defined as the difference between the temperatures at which the HS state is 20 and 80% occupied (ΔT_{60}, Figure S4.1 in the SI).[6,7] The abruptness of the transition is strongly related to the notion of cooperativity, that represents the capability of one SCO centre to directly influence the spin state of a neighbouring one within the material.

An abrupt spin crossover transition is often a first step toward hysteretic behaviours. A commonly adopted picture links this abruptness to the strength of intermolecular interactions 'cooperativity' whereby a spin transition in one molecule promotes a

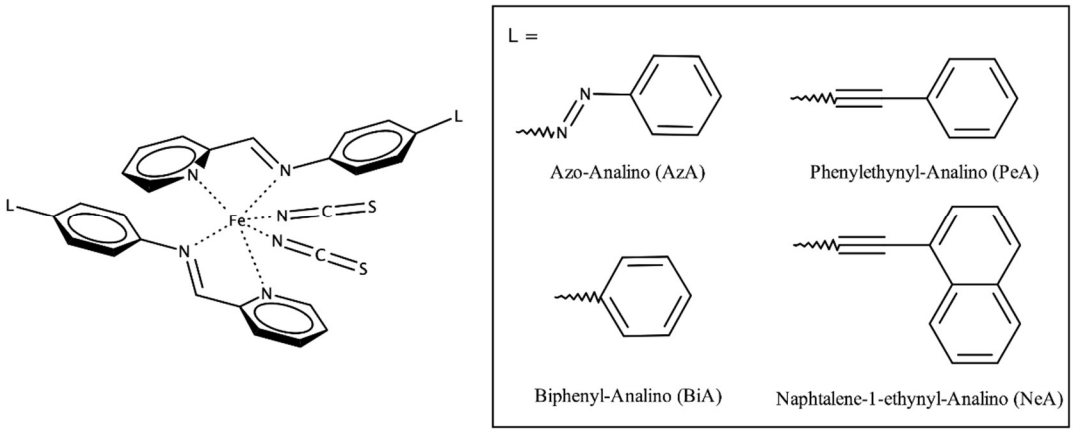
transition in a neighbouring molecule, leading to a sharp spin crossover transition. [6,8-14] The nature of a transition thus depends not only on the identity of a complex, but also on its crystal structure.[15] The nature of the intermolecular interactions can strongly differ from one compound to another and several of them can coexist at the same time. For instance, it has been suggested that subtle interactions such as H···H contacts could play a crucial role in SCO phase transitions, even when they are partially repulsive.[16] Recent work has applied Hirshfeld surface analysis to obtain an overview of all intermolecular contacts in SCO crystals, which provides a description of the intermolecular topology of SCO structures, especially where a single commonly identifiable or prominent interaction cannot be established.[17] This is particularly true in the case of polymorphs.

The complexity of the intermolecular interaction topology requires a more complete description. Not only for a better view of this network but also to give new hints to chemists for the synthesis toward a better control of this interacting network. This chapter tackles this challenge, focusing on a specific SCO family exhibiting polymorphism and various types of intermolecular interaction and which has already been used to establish structure - SCO properties relationships: the [Fe(PM-L)$_2$(NCS)$_2$] family of complexes (Scheme 4.1).

In this family, transition abruptness ΔT_{60} can vary between 5 and 97 K. Moreover, differing behaviour is seen for polymorphs as for example for [Fe(PM-BiA)$_2$(NCS)$_2$] (PM-BiA = N-(2-pyridylmethylene)-4-aminobiphenyl).[7,18] The abruptness of transitions has been associated with a short S…H-C intermolecular contact in the crystal structures of the HS forms (Figure *4.1*.i): where this contact is short the spin transition tends to be sharp [6] (Figure *4.1* .ii), suggesting that the strength of this interaction leads to increased cooperativity. While the trend is applicable to many members of the PM-L family generally, it does not apply to Fe(PM-NeA)$_2$(NCS)$_2$ (NeA = naphthalene-1-ethynyl-anilino), which has short S…H-C contacts (S…C = 3.438 Å) but a very gradual transition (ΔT_{60} = 97 K).[19]

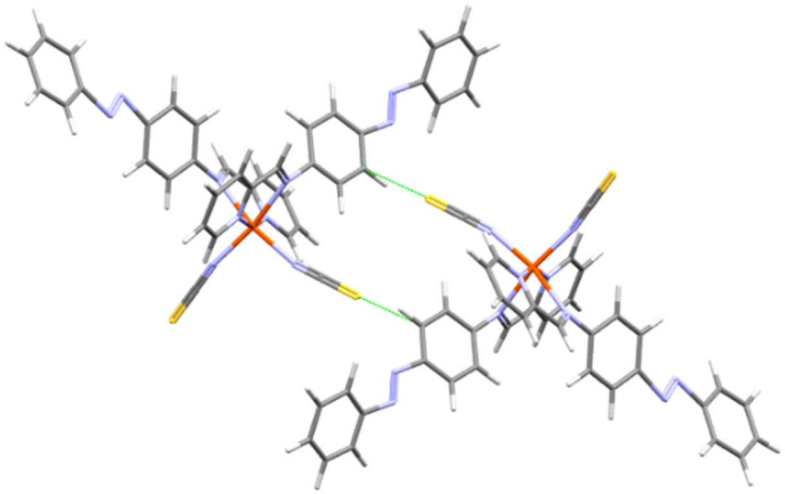
In this chapter, intermolecular packing energy calculations are applied to the crystal structures of the Fe(PM-L)$_2$(NCS)$_2$ family of spin crossover complexes with the aim

of obtaining a more systematic overview of the energies of the intermolecular interactions which define the cooperativity pathways in these structures. The calculations have been performed with the PIXEL method,[20] in which molecule-molecule (as opposed to atom-atom) energies are calculated semi-empirically using *ab initio* molecular electron densities. The calculations yield not only a total lattice energy, but also its breakdown into individual molecule-molecule contributions. Each molecule-molecule energy is further decomposed into electrostatic, polarisation, dispersion and Pauli repulsion terms, providing insight not only on the strength, but also on the physical nature of the intermolecular interactions.



Scheme 4.1: $Fe(PM\text{-}L)_2(NCS)_2$ Complex. PM-L Ligands for studied structures.

i.



ii.

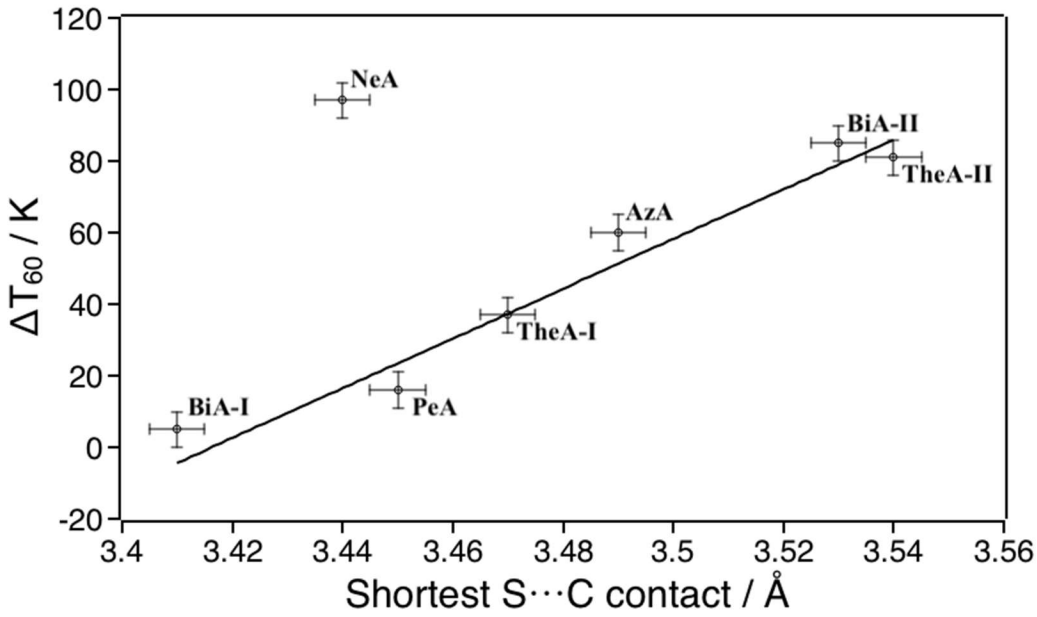


Figure 4.1: i. Shortest S…H-C (measured as S...C) interaction in $Fe(PM\text{-}AzA)_2(NCS)_2$. ii. Relationship between C...S interaction distance, in the HS structures and SCO width in $Fe(PM\text{-}L)_2(NCS)_2$ in Scheme 1. Two points are plotted for PeA which has slightly different values of ΔT_{60} in the forward and reverse transitions between spin states. BiA-I and BiA-II correspond to different polymorphs. The values plotted are available in Table S4.3.

4.2. Methodology

4.2.1. Structural Data

Four spin crossover complexes were investigated (Scheme 1). Of these, the PM−BiA derivative exists in two polymorphic forms.

[Fe(PM-AzA)$_2$(NCS)$_2$] has a gradual SCO transition with ΔT_{60} = 61 K. Structures are available on the Cambridge Structural Database (CSD) for the HS form at 290 K (CSD Refcode: XECNAU35) and the LS form at 110 K (XECNAU07). The orthorhombic form-I of [Fe(PM-BiA)$_2$(NCS)$_2$] (HS, 290 K, RONPIT01; LS 140 K, RONPIT02) has an abrupt spin transition (ΔT_{60} = 5 K),[21] while the monoclinic phase-II (HS, 290 K RONPIT04, LS 120 K RONPIT05) has a much more gradual transition (ΔT_{60} = 81 K).[7] The temperature induced spin−transition of [Fe(PM−PeA)$_2$(NCS)$_2$] is accompanied by a phase transition upon heating/cooling from LS *Pccn* (140 K, NOWBIK) to HS $P2_1/c$ (290 K, NOWBIK01) with a relatively small ΔT_{60} of 30 K in the HS -> LS direction and 37 K for LS -> HS.[22] The average of these values (33.5 K) was used in the analysis below. [Fe(PM-NeA)$_2$(NCS)$_2$] has the most gradual temperature-induced SCO transition found in the PM-L family (ΔT_{60} = 97 K). Structures for both HS (290 K, COMQUR) and LS (120 K, COMQUR01) forms are available.[19] As all complexes do not crystalize with the same lattice symmetry, equivalent asymmetric units for HS and LS structures were obtained in $P2_1/c$ space group with the aid of the EQUIVSTRU utility on the Bilbao Crystallographic Server.[23-25] The coordinates used for the calculations are available in the SI.

4.2.2. PIXEL calculations

Lattice energies and intermolecular interaction energies were calculated using the semi-empirical computational technique PIXEL.[20,26,27] PIXEL calculates energies by modelling each molecular component as a grid of small cubes (pixels) of electron density. Interactions are calculated between a central reference molecule and other molecules within a cluster generated from the crystallographic space group symmetry. Intermolecular energies are calculated from the sum of electrostatic, polarisation, dispersion and (Pauli) repulsion terms accumulated from each pixel-pixel combination in a dimer. The sum of all cluster interaction energies gives the lattice energy. In this study the cluster radius was 20 Å, and the molecular electron densities were obtained

in steps of 0.08 Å from GAUSSIAN-09 with the 6-31G** basis set at the B3LYP level of theory.[28] The PIXEL calculations themselves were accomplished with the CLP-PIXEL suite within the MrPIXEL interface[29] using a condensation level of 4 (i.e. the original pixels from GAUSSIAN were combined into 4×4×4 blocks of dimension 0.32 Å).[30] The parameters applied to the transition metals were those derived by Maloney *et al.*[31] C−H distances were reset to 1.089 Å to correct for the systematic shortening of bonds involving hydrogen in X-ray crystal structures.

4.2.3. Visualisation of results

Intermolecular interactions were visualised in Mercury[32] using energy frameworks, originally devised by Spackman and co-workers for implementation in the CrystalExplorer software.[33,34] The width and colour of the struts drawn between molecules represent the intermolecular energy. In the Figures below green struts represent stabilising interactions (energy < 0), while pink struts define destabilising interactions (energy > 0). The thickness represents the strength of the interaction.

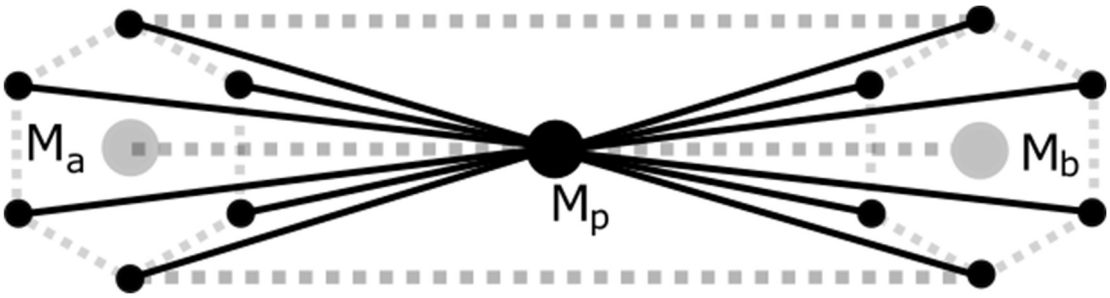


Figure 4.2: Example of strut construction showing new atoms placed between Molecule M_a and M_b and bonds between newly placed atoms. Magnitude of hexagons dictated by energy over value n where E/n produces a sensible strut size.

The frameworks were generated using the CSD Python API (application programming interface). A vector M_a-M_b defined between the molecular centroids of each interaction identified in a PIXEL calculation (Figure 4.2). An atom (radium or meitnerium for stabilising and destabilising contacts, respectively) is placed at the mid-point of the vector, M_p. Six deuterium atoms are placed orthogonal to the vector M_a-M_b and at a distance E/n from both M_a and M_b. E is the energy of the interaction in kJ mol^{-1} and n determines the diameter of the struts of the framework. Direct visual comparisons between frameworks can be made provided the same value of n is used; in this work n was chosen to be 200. The struts are then constructed by representing

the atom at M_p as a hexagonal prism with the vertices at the deuterium atoms. The updated structure is saved as a MOL2 format file, which can then be visualised in Mercury. As has been described by Spackman, frameworks may be generated for each separate component of the interaction energy (electrostatic, polarisation, dispersion, repulsion) or as the total.

In addition, comparisons can be made between isostructural crystal structures, or those where interactions may be mapped from one structure to another, by calculating 'difference frameworks' where the strut sizes are related to the energy difference (ΔE) between the interactions in each structure. In this case a value of n = 25 has been used. The scripts for producing these plots are available from www.crystal.chem.ed.ac.uk.

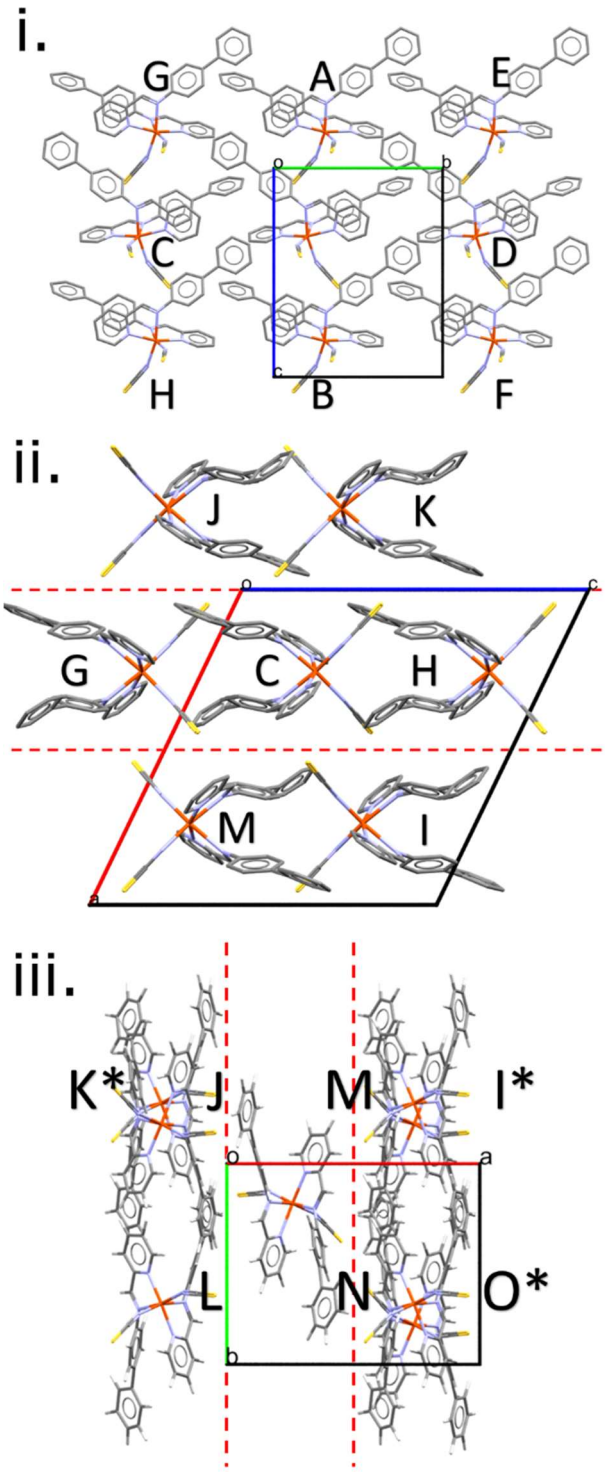


Figure 4.3: Stacking layers in $Fe(PM\text{-}BiA)_2(NCS)_2$ polymorph-II with layers (separated by red−hashed lines) parallel to the layer−axis, viewed along a (**i.**), b(**ii.**) and c(**iii.**) axes respectively. reference molecule is unlabelled and in the centre of each perspective (in the background overlapped with C in (ii)). First coordination sphere contacts to initial molecule are labelled beside molecule. Where molecules are in background, labels are denoted with *.

4.3. Results

4.3.1. Crystal Packing in Fe(PM−L)$_2$(NCS)$_2$ Structures

The crystal structures in the Fe(PM-L)$_2$(NCS)$_2$ family can be described in terms of the formation of layers in which the molecules are arranged so that a molecular axis drawn between the metal atom and the molecular centroid is parallel to the layer (Table S4.1).[35] The unit cell contains two offset layers in which the molecular axes point in opposite directions leading to an alternating stacking sequence. The closest layer spacing (−0.78 Å) occurs for the HS form of Fe(PM-NeA)$_2$(NCS)$_2$, the negative sign indicating that there is some interpenetration between the layers (Figure S4.3). The largest (1.4 Å) occurs in Fe(PM-BiA)$_2$(NCS)$_2$-I LS (Figure S4.4).

The general features of packing and the pattern of intermolecular interactions in this family of spin crossover complexes can be illustrated using Fe(PM-BiA)$_2$(NCS)$_2$ in polymorph-II of its high spin form (CSD Refcode RONPIT04, Figure 4.3). Within a layer, each molecule makes energetically significant contacts with eight other molecules, labelled A-H in 4.3 .i and shown individually with their energy breakdowns in Table S4.2 in the SI. The shortest centroid-centroid distances (8.719 Å) occur along chains which run parallel to the *c*-axis, interactions A and B in *4.3* .i. The molecule-molecule energy of these interactions is −97.9 kJ mol^{-1}, and the shortest atom-atom contacts formed involve the thiocyanate sulfur atoms in one molecule and hydrogen atoms in the next (S1…H19, 2.86 Å). These interactions have large electrostatic and dispersion components. Adjacent chains in the same layer related by lattice translations along **b** interact more weakly (C/D, −20.9 kJ mol^{-1}). The inter-chain contacts formed diagonally to molecules E/F and G/H in Figure 4.3 .i have energies of −14.6 kJ mol^{-1} and −22.6 kJ mol^{-1}, respectively. The largest contributing energy term in the inter-chain interactions (C to H contacts in figure 4.3.i) is dispersion.

In Fe(PM-BiA)$_2$(NCS)$_2$ polymorph-II there are seven interlayer interactions (Figure 4.3 .ii), in other members of the family (such as the complexes with PM-AzA, PM-PeA and PM-BiA polymorph-I) there are eight, giving total molecular coordination numbers of 15 or 16. The seven contacts in Fe(PM-BiA)$_2$(NCS)$_2$-II (labelled I-O) vary in energy between −80.7 and −27.4 kJ mol^{-1}. The strongest interactions, I and J, are formed across inversion centres and involve electrostatic thiocyanate…H and

dispersion-dominated π…π interactions, respectively. In other contacts, the electrostatic term is largest where the closest atom-atom contacts involve the thiocyanate ligands; where the contact is between rings, dispersion dominates. The shortest C…S contact (part of interaction denoted O), which has been implicated in controlling the SCO transition width in previous work (see above), forms diagonally between molecules in adjacent layers (as shown in Figure 4.3 .iii) for all structures, with thiocyanate group in one molecule pointing directly towards the phenyl H-atoms in the other.

Further inter-layer interactions occur where the thiocyanate is orientated towards an aromatic ring, creating an S…aromatic ring contact that is dominated by the electrostatic term. The strength of this interaction varies considerably between structures and spin-states (see Section 4.3.2), with the C…S distance varying between 3.41 and 3.55 Å. The correlations involving C…S interactions developed in previous work have been based on the distance measured in the crystal structure of the complex in its HS configuration. As these distances change across a SCO transition their influence on cooperativity also changes, so that the interactions which govern cooperativity in the HS to LS transition may be different from those in the LS to HS transition.

It is thought that strong intermolecular contacts, of the type described above, are the source of cooperativity in spin crossover transitions.[12,36] The following sections aim to show how the overall view of intermolecular interaction energies provided by the PIXEL results can help to pin-point the structural features which determine transition abruptness in the $Fe(PM\text{-}L)_2(NCS)_2$ family of complexes.

4.3.2. Relating SCO behaviour to structural parameters and PIXEL energies

The shortest C…S Interaction

The suggestion that spin transition abruptness is related to the strength of the intermolecular interactions containing the short C…S contacts may be analysed in terms of interaction energies using the PIXEL results for interaction O (see Figure 4.3 .iii) and its analogues in other structures (see also Table S4.3). While there are substantial differences in intermolecular interaction energies in these systems, no simple trend can be identified between the transition abruptness ΔT_{60} and either the

total dimer energy in the high or low spin forms or the change in dimer energy between spin states (Figure 4.4 .i). The results suggest that SCO behaviour cannot be fully described or predicted from the energy of the C…S interaction, and therefore the investigation moves to focus on whether parameters such as lattice energy or inter-layer spacing, which are features of the whole crystal structure, might be more effective.

Lattice Energies

The lattice energies calculated using PIXEL are available for each structure and are listed in Table S4.4. The lattice energy change between spin-states at 110 and 290 K is in the order of –20 to –50 kJ mol^{-1} for all complexes, with the two polymorphs of $Fe(PM\text{-}BiA)_2(NCS)_2$ having very similar lattice energies for both spin states (E_L(HS) and E_L(LS), respectively). This may explain why both polymorphs are observed under ambient conditions, though the energies are too similar to state with confidence which is the more stable form. From these results there is no clear correlation between transition width and the lattice energies of either the HS or LS structures. Neither is there a correlation with change in lattice energy between spin-states for the HS → LS transition, defined as $\Delta E_L = E_L(LS) - E_L(HS)$ (Figure 4.4 .ii).

Layer Spacing

Previous studies find no clear trend relating to the isotropic cell contraction (ΔV_{SC}) and the SCO characteristics, but do point towards the anisotropy of the cell contraction as a parameter relating to the abruptness of transition.[6] The change in layer spacing is a component of the anisotropic cell volume contraction and thus provides information on the nature of crystal packing changes in relation to SCO behaviour for $Fe(PM\text{-}L)_2(NCS)_2$ with this layered packing. Although there is no apparent trend between the level of interpenetration of the layers described above and the abruptness of transition, the change in layer separation between spin-states does suggest that large negative changes in layer separation correlate with sharper SCO (Figure 4.4 .iii), though this correlation does not extend to the broader transitions.

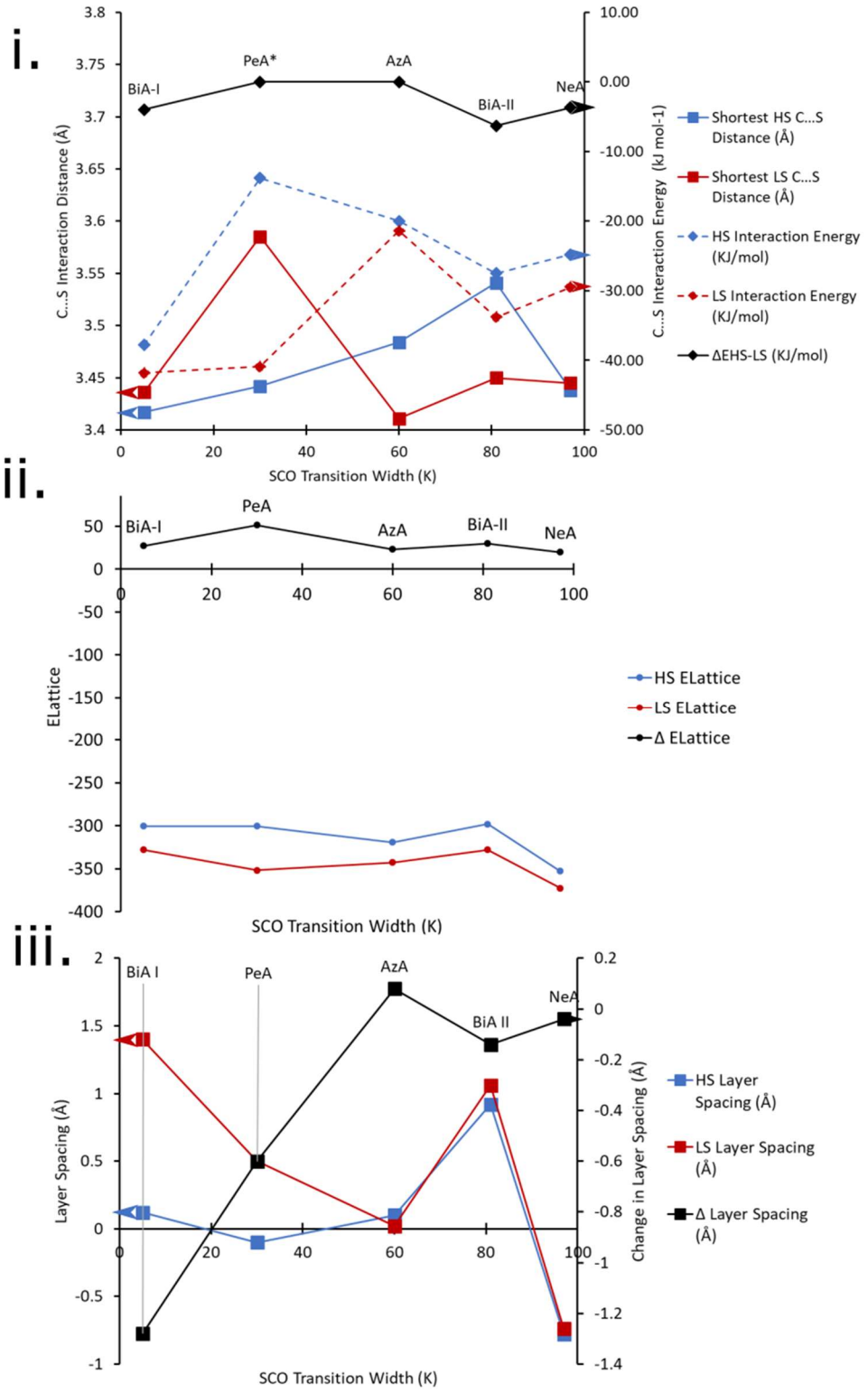


Figure 4.4: i. SCO transition width against the shortest C..S interaction distances and PIXEL interaction energies. Data for both HS and LS forms are plotted. ii. SCO transition width against HS/LS PIXEL lattice energies. iii. SCO transition width against HS and LS layer spacing.

Energy Frameworks

Energy frameworks were devised by Spackman and co-workers as a way of rapidly visualising intermolecular interactions in a crystal structure.[34] The framework is constructed by linking pairs of molecules with struts, where the width of a strut is proportional to the intermolecular energy: thick struts correspond to strong interactions. Energy frameworks have been shown to be helpful in explaining the source of physical properties such as the role of π-π stacking and hydrogen bonding in 1-hydroxypyrene crystal morphology, and in exploring polymorphism in small organic molecules such as glycine.[33,37]

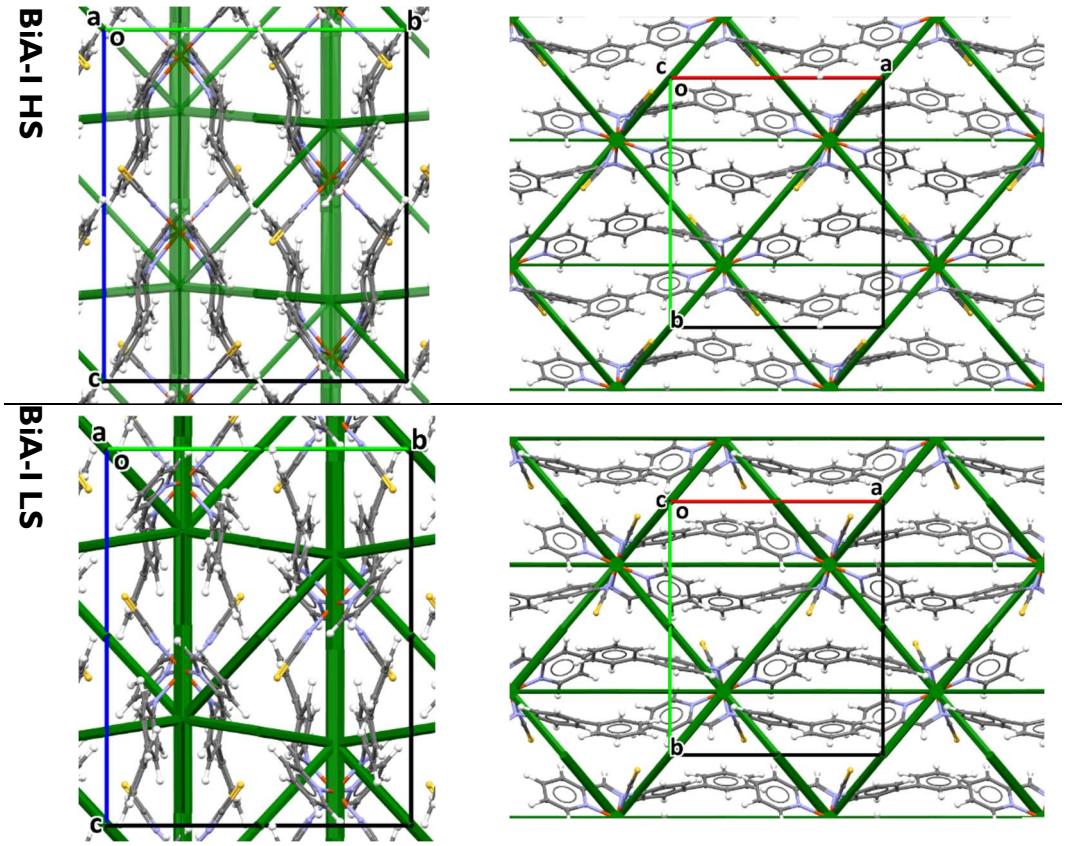
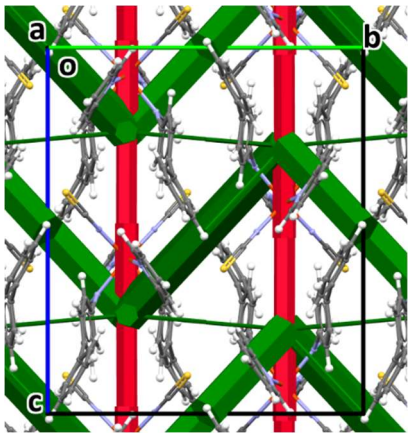


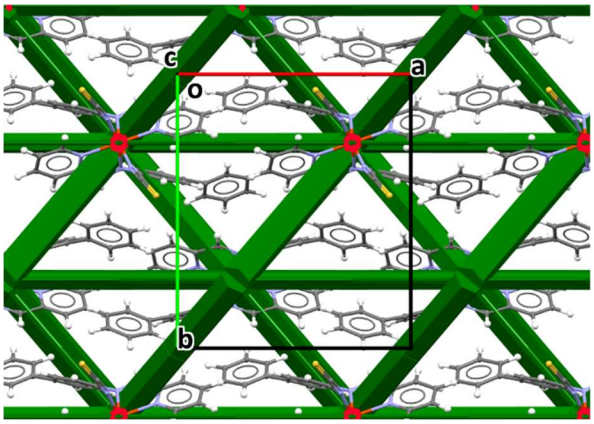
Figure 4.5: Comparison of HS (top) & LS (bottom) energy frameworks for Fe(PM-BiA)$_2$(NCS)$_2$ polymorph I viewed along a axis (left) and c axis (right).

The differences between the intermolecular interactions in the high and low spin states are not obvious from a simple comparison of their energy frameworks, as shown for the PM-BiA-I structures in Figure 4.5. As an alternative it is possible to produce an energy "difference" framework in which each strut represents a dimer interaction with a width proportional to its change in energy *E* across the HS → LS transition, $\Delta E = E(\text{LS}) - E(\text{HS})$.

BiA-I

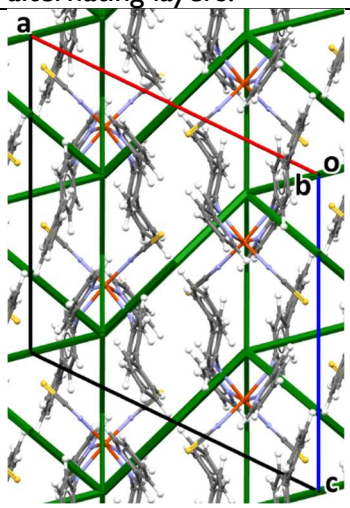


HS−>LS interaction energy changes viewed along **a** axis between alternating layers.

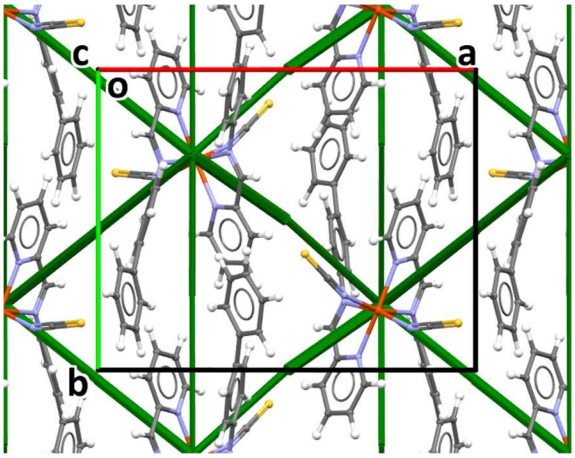


HS−>LS interaction energy changes viewed along **c** axis down multiple layers.

BiA-II



HS−>LS interaction energy changes viewed along **b** axis between alternating layers.



HS−>LS interaction energy changes viewed along **c** axis down multiple layers.

Figure 4.6: Energy difference frameworks for polymorphs of $Fe(PM\text{-}BiA)_2(NCS)_2$. For clarity, struts are only shown for the intermolecular first molecular coordination sphere (i.e. first nearest neighbours) where the interaction energy changes by more than 2.5 kJ mol^{-1}. The same criteria apply to the other difference frameworks shown in this chapter.

The difference frameworks for the two polymorphs of $Fe(PM\text{-}BiA)_2(NCS)_2$ are shown in Figure 4.6. Green struts correspond to interactions where the energy becomes more stabilising (i.e. $\Delta E < 0$) during the HS to LS transition. Pink struts show interactions which are destabilised (i.e. $\Delta E > 0$). The magnitude of the energy change is shown by the thickness of the strut, as usual. The difference framework of BiA-I, which has a very sharp transition, is characterised by thick stabilising (green) and destabilising (pink) struts. In this structure interactions C, D, M and N are stabilised by over -10 kJ mol^{-1} over the course of the HS→LS transition, while interactions A and B, consisting of close S..H contacts, are destabilised (see Tables S4.5 and S4.6 in the SI). These changes are largely electrostatic in origin and are a result of the reorientations and changes in S…H and S…aromatic ring intermolecular contact distances. By contrast the difference framework of BiA-II, which has a broad SCO transition, has much thinner stabilising struts and no destabilising changes.

A similar pattern emerges in other structures. Figure 4.7 .i shows the energy difference framework for $Fe(PM\text{-}AzA)_2(NCS)_2$ in which the generally thin green struts and the absence of pink struts correctly suggests that it should, like $Fe(PM\text{-}BiA)_2(NCS)_2$ form II, have a broad SCO transition.

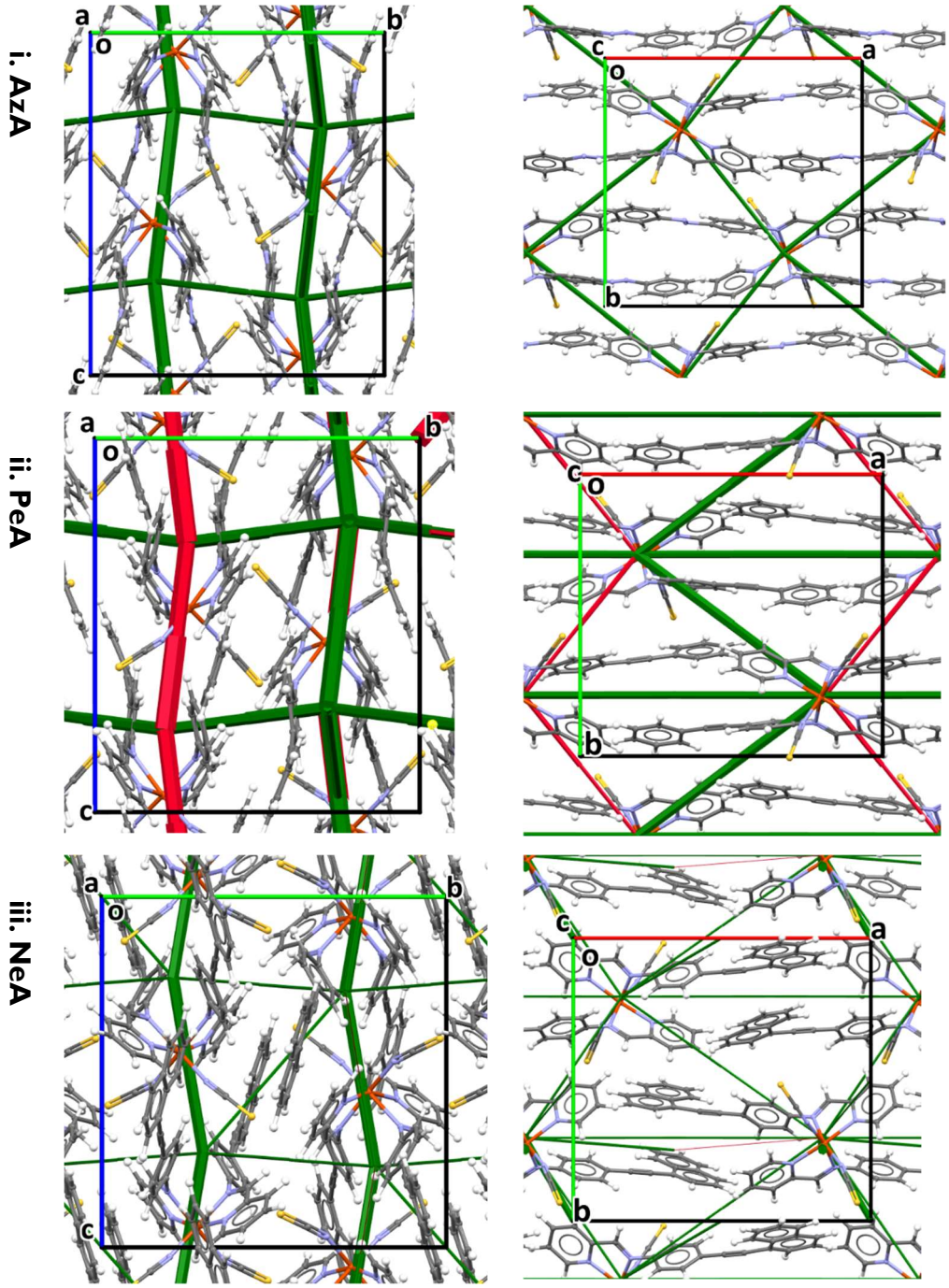


Figure 4.7 i. Energy difference framework for $Fe(PM\text{-}AzA)_2(NCS)_2$ along a axis and c axis respectively. ii. Energy difference framework for $Fe(PM\text{-}PeA)_2(NCS)_2$ along c axis and a axis respectively. iii. Energy difference framework for $Fe(PM\text{-}NeA)_2(NCS)_2$ along a axis and c axis respectively.

The energy differences framework for $Fe(PM\text{-}PeA)_2(NCS)_2$ shows much larger energy changes between spin-states, with both stabilizing and destabilizing changes (Figure 4.7 .ii), though these are less than in the transition for $Fe(PM\text{-}BiA)_2(NCS)_2$-I. The

transition is therefore expected to be more abrupt than for the AzA complex, but less so than for BiA-I. The $Fe(PM\text{-}NeA)_2(NCS)_2$ complex undergoes the most gradual transition (97 K) in the PM-L family. As expected, the energy difference framework for this complex demonstrates much smaller energy changes between spin states, as shown in Figure 4.7 .iii.

The presence of prominent stabilising and destabilising energy changes over the HS → LS transition is seen to be associated with a sharp SCO transition, while smaller changes and the absence of significant destabilising changes are seen to be associated with broad spin crossover transition. The relationship can be quantified in a plot of the sum of the magnitudes of all the interaction energy changes ($\sum|E_{Tot}|$ = the sum of the widths of the struts shown in Figure *4.6* and Figure *4.7*) against ΔT_{60}, which is linear (Figure 4.8) with a correlation coefficient (*r*) equal to −0.94. The correlation coefficient is negative because ΔT_{60} decreases as the energy changes increase. It is worth noting that this linear trend is observed on the whole PM-L family, including the$[Fe(PM\text{-}NeA)_2(NCS)_2]$ complex that did not follow the previous general trend based on the HS S…H contact distance. It underlines the relevance of exploring the whole interaction map instead, as facilitated by the energy frameworks approach.

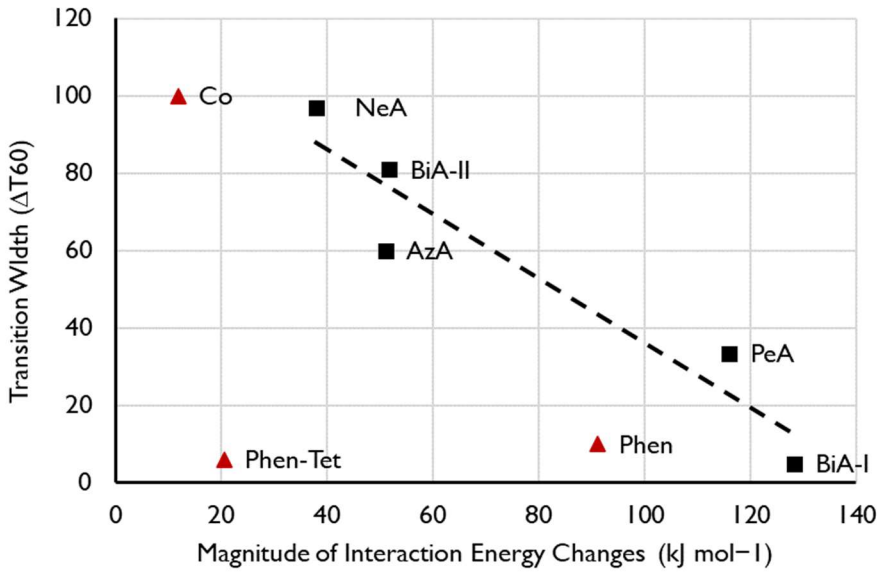


Figure 4.8: The transition width (ΔT_{60}) plotted against the sum of the of absolute total interaction energy changes for $Fe(PM\text{-}L)_2(NCS)_2$ SCO complexes (black squares) and three other SCO complexes (red trianges, see Section 4.4). The line is fitted against the PM-L family of structure only, and has an equation of $y = -0.8339x + 119.54$.

Linear trends are also seen for the most positive and most negative total interaction energy changes for each structure (Table S4.7), corresponding to the thickest pink and green struts in the energy frameworks, and even to the magnitudes of the contributing energy terms themselves (Figure S4.5), such as the sum of electrostatic energy magnitudes ($\sum_i \left|E_{i(Elec)}\right|, r = -0.84$), the most positive total energy changes ($Max\ E_{Tot}, r = -0.92$), and the most stabilizing total energy change ($Min\ E_{Tot}, r = 0.93$). The most linear correlation (r = −0.97) is seen for the most positive electrostatic energy change. These correlations are much stronger than those found for the sum of the signed values (i.e. where signs of changes are taken into account, r = 0.42). This is consistent with the absence of a correlation between ΔT_{60} and the change in lattice energy (Figure 4.4 .ii).

The difference seen for the analysis of magnitudes and sums of energies may seem paradoxical, but abrupt SCO transitions are associated with the presence of thick struts in the energy frameworks which can correspond to large positive or negative energy changes. When signed energies are summed, large positive and negative energy changes can cancel each other out and their influence is lost in the total energy. The implication is that an abrupt transition is one in which the intermolecular interactions are able to accommodate the strain generated by the change in volume of the SCO complex in a flexible way, which may change individual terms substantially, but does not necessarily incur a large overall change in total energy.

4.4. Application beyond the Fe(PM-L)$_2$(NCS)$_2$ Family

The results discussed so far have focussed on a single family of SCO complexes. Although the complexes are not strictly isostructural, the crystal packing and intermolecular interactions are consistent across the series, allowing direct comparisons of interactions. The quantitative correlation shown in Figure 4.8 would not be expected to extend generally, beyond the PM-L family, because different classes of SCO material will differ in polarity, molecular size and contacts, so that the scale of the intermolecular interaction energy changes will also be different.

The difference frameworks for three complexes from beyond the PM-L family are shown in Figure 4.9, with energy data available in Table S4.8. The complex [Fe(Phen)$_2$(NCS)$_2$] ('Phen', CSD Refcodes HS:KEKVIF, LS:KEKVIF01) is similar to the PM-L family in having thiocyanate ligands, but the phenanthroline ligands are smaller and better suited for graphitic stacking. The difference framework (Figure 4.9 .i) consists of prominent green and pink struts which are consistent with its abrupt SCO transition (ΔT_{60} = 10 K), but in Figure 4.8 the point for this complex lies close to, but not on, the PM-L correlation line.

In the cobalt complex PUYROS ('Co', HS: PUYROS01, LS: PUYROS) the metal binds to aromatic imines which are similar in some respects to PM-L ligands, but thiocyanates are absent. In common with most Co(+2) complexes its SCO transition is very broad (ΔT_{60} is quoted as > 100 K, for the purposes of plotting ΔT_{60} = 100 K). Accordingly, its difference framework (Figure 4.9 .ii) is virtually featureless.

By contrast, Fe(Tet-Phen)$_2$(NCS)$_2$ ('Phen-Tet', HS:QIDJET, LS:QIDJET01), which contains phenanthroline ligands substituted with anionic tetrazolyl groups, has a very abrupt transition (ΔT_{60} = 6 K), again in line with the prominent green and pink struts found in its difference framework, but it does not fit the PM-L correlation at all (Figure 4.8).

The value of ΔT_{60} for QIDJET is similar to that of Fe(PM-BiA)$_2$(NCS)$_2$ form I, but its framework (Figure 4.9 .iii) lacks the prominent green and pink struts which characterise the plot for PM-BiA complex in Figure 4.6 .i. Nevertheless, what the two plots do have in common is that both contain stabilising green and destabilising

pink struts of similar width. This observation is consistent with the trends seen in the PM-L family, in which prominent struts of both types are present in the fast-transitioning complexes. The plot for $Fe(PM\text{-}AzA)_2(NCS)_2$ may contain prominent green struts, but the pink struts are much thinner, and the SCO transition is relatively broad.

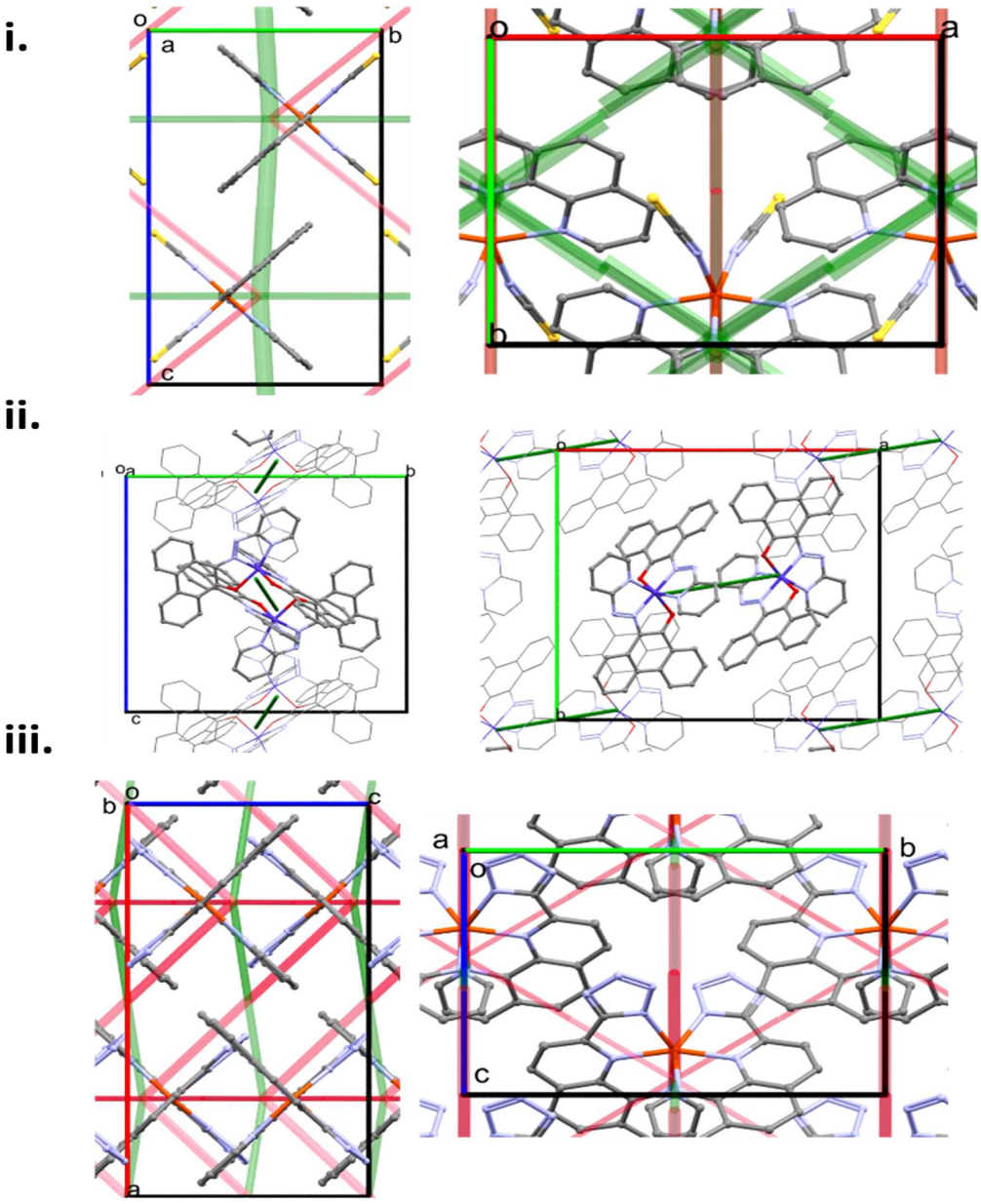


Figure 4.9: Energy difference frameworks for i. $Fe(phen)_2(NCS)_2$, ii. bis(10-((pyridine-2-yl)diazenyl)phenanthrene-9-olato)-cobalt and iii. $Fe(phen\text{-}Tetrazol)_2$. All metals are in the +2 oxidation state.

It seems that an abrupt transition requires both prominent stabilising and destabilising changes to be present. This follows the conclusion that an abrupt transition needs intermolecular interactions that are able to accommodate the strain generated by the change in volume of the SCO complex. Although the scale of these changes varies with the ligands, a plot of ΔT_{60} against the range of energy changes (Range ΔE_{Tot}) normalised by dividing by the magnitude of the maximum interaction energy change (i.e. the thickest strut) (Figure 4.10) shows that fast and slow transitioning complexes cluster in two different regions.

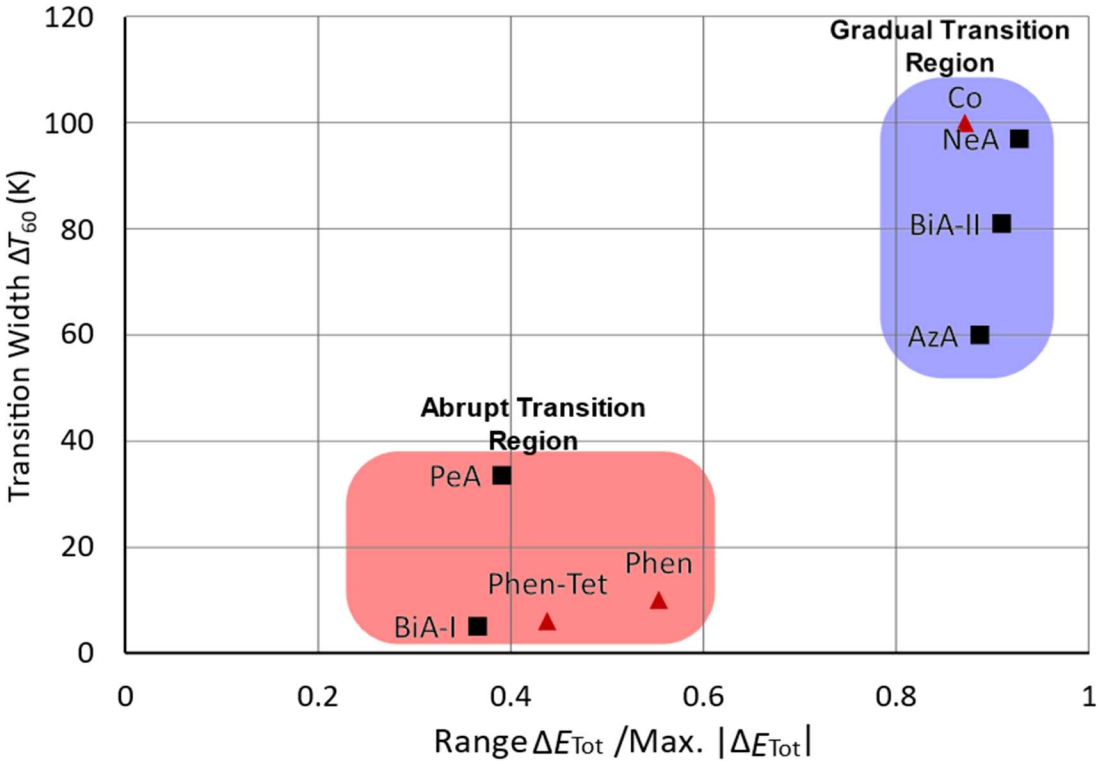


Figure 4.10: Transition width (ΔT_{60}) against the quotient of the difference between largest destabilizing and stabilizing changes (Range ΔE_{Tot} /max $|\Delta E_{Tot}|$) for $Fe(PM\text{-}L)_2(NCS)_2$ SCO complexes (black squares) and other SCO complexes (red trianges). Range ΔE_{Tot} is defined as |most +ve ΔE_{Tot}|−|most −ve ΔE_{Tot}|.

Such a normalisation fully addresses the non-fitting compounds in the SCO abruptness versus energy changes representation leading to two distinct regions in the diagram (Figure 4.10) showing that this approach based on global evaluation of intermolecular interactions can be applied to numerous SCO compounds even if they do not belong to the same family.

4.5. Conclusions

The aim of this work has been to explore the relationship between the abruptness of spin-crossover transitions and intermolecular interactions. Only crystalline solids have been considered. The study has concentrated on the PM-L family of Fe(+2) complexes. The packing is consistent within this series of crystal structures, enabling direct comparisons to be made between different members of the family. A full set of structural and magnetic data is also available for both HS and LS forms of all complexes. The crystal structures can also be described with one molecule in the asymmetric unit, making them suitable for lattice energy calculations using PIXEL.

No correlation was found between abruptness and lattice energy, the intermolecular energy involving shortest HS C…S contact, or the changes in layer stacking which occur across an SCO transition. A more consistent trend emerges by considering the changes in individual intermolecular interaction energies, and there is a linear variation in ΔT_{60} with the sum of the magnitudes of the interaction energy changes within the first molecular coordination sphere (Figure 4.8). These changes can be visualised in energy difference frameworks, which could also be considered as illustrating the cooperativity pathways of an SCO transition.

Abrupt spin-crossover transitions require some large stabilising and destabilising changes in intermolecular energies. It does not seem to matter which interactions are involved: the largest positive and negative energy changes depend on the structure, and, despite the packing similarities across the PM-L family, there is no consistency in which interaction shows the largest energy change.

The quantitative trend established for the PM-L family does not directly extend to other classes of SCO complexes, and it is suspected that this is because the scale of the energy changes which occur are strongly dependent on the ligands present. Instead, it is necessary to assess the significance of individual contact energy changes in the context of the overall magnitude of the changes taking place. A clear correlation was found, independent of the chemical family, when normalized interaction energy changes were used against the abruptness of the transition. Fast and slow transitioning complexes cluster together in different regions of a plot of ΔT_{60} against Range ΔE_{Tot} /max $|\Delta E_{Tot}|$ (Figure 4.10).

These results suggest that the abruptness of an SCO transition is related to the accommodation of strain which is generated as the volumes of individual molecules change with spin state. Large changes could indicate that strain is being accommodated in specific, flexible interactions instead of needing to be propagated through the entire structure, destabilising changes being compensated for by stabilising changes. For gradual transitions, the small energy changes and lack of compensating energy changes appear to cause a slower propagation of spin transitions through the system.

4.6. References

(1) Goodwin, H. A. Spin Transitions in six-coordinate iron(II) complexes. *Coordination Chemistry Reviews* **1976,** *18* (3), 293.

(2) Gütlich, P., Berlin, Heidelberg, 1981; p 83.

(3) Gütlich, P.; Gaspar, A. B.; Garcia, Y. Spin state switching in iron coordination compounds. *Beilstein Journal of Organic Chemistry* **2013,** *9*, 342.

(4) Gütlich, P.; Link, R.; Trautwein, A. *Mössbauer spectroscopy and transition metal chemistry*; Springer-Verlag, 1978.

(5) Wolny, J. A.; Schünemann, V.; Németh, Z.; Vankó, G. Spectroscopic techniques to characterize the spin state: Vibrational, optical, Mössbauer, NMR, and X-ray spectroscopy. *Comptes Rendus Chimie* **2018,** *21* (12), 1152.

(6) Guionneau, P.; Marchivie, M.; Bravic, G.; Letard, J.-F.; Chasseau, D. Structural Aspects of Spin Crossover — Example of the [FeIILn(NCS)2] Complexes. *ChemInform* **2005,** *36* (47).

(7) Marchivie, M.; Guionneau, P.; Letard, J.-F.; Chasseau, D. Towards direct correlations between spin-crossover and structural features in iron(II) complexes. *Acta Crystallographica Section B* **2003,** *59* (4), 479.

(8) Köhler, C. P.; Jakobi, R.; Meissner, E.; Wiehl, L.; Spiering, H.; Gütlich, P. Nature of the phase transition in spin crossover compounds. *Journal of Physics and Chemistry of Solids* **1990,** *51* (3), 239.

(9) Nemec, I.; Herchel, R.; Trávníček, Z. The relationship between the strength of hydrogen bonding and spin crossover behaviour in a series of iron(iii) Schiff base complexes. *Dalton Transactions* **2015,** *44* (10), 4474.

(10) Guionneau, P. Crystallography and spin-crossover. A view of breathing materials. *Dalton Transactions* **2014,** *43* (2), 382.

(11) Halcrow, M. A. Structure: Function Relationships in Molecular Spin-Crossover Materials. *Spin-Crossover Materials: Properties and Applications* **2013**, DOI:10.1002/9781118519301.ch5 10.1002/9781118519301.ch5, 147.

(12) Spiering, H. In *Spin Crossover in Transition Metal Compounds III*; Springer Berlin Heidelberg: Berlin, Heidelberg, 2004, DOI:10.1007/b95427 10.1007/b95427.

(13) Haddad, M. S.; Lynch, M. W.; Federer, W. D.; Hendrickson, D. N. Spin-crossover ferric complexes: curiosities observed for unperturbed solids. *Inorganic Chemistry* **1981,** *20* (1), 123.

(14) Kambara, T. Theory of high-spin⇄low-spin transitions in transition metal compounds induced by cooperative molecular distortions and lattice strains. *The Journal of Chemical Physics* **1981,** *74* (8), 4557.

(15) Collet, E.; Guionneau, P. Structural analysis of spin-crossover materials: From molecules to materials. *Comptes Rendus Chimie* **2018,** *21* (12), 1133.

(16) Tailleur, E.; Marchivie, M.; Negrier, P.; Denux, D.; Massip, S.; Mondieig, D.; Chastanet, G.; Guionneau, P. Using polymorphism to master the spin crossover mechanism in [Fe(PM-PeA)2(NCSe)2]. *CrystEngComm* **2019,** *21* (41), 6246.

(17) Guionneau, P. Crystallography and spin-crossover. A view of breathing materials. *Dalton transactions (Cambridge, England : 2003)* **2013,** *43*.

(18) Tao, J.; Wei, R.-J.; Huang, R.-B.; Zheng, L.-S. Polymorphism in spin-crossover systems. *Chemical Society Reviews* **2012,** *41* (2), 703.

(19) Létard, J.-F.; Kollmansberger, M.; Carbonera, C.; Marchivie, M.; Guionneau, P. Structural, magnetic and photomagnetic study of the [Fe(PM–NEA)2(NCS)2] spin crossover complex. *Comptes Rendus Chimie* **2008,** *11* (10), 1155.

(20) Gavezzotti, A. Calculation of lattice energies of organic crystals: The PIXEL integration method in comparison with more traditional methods. *Z. Krist.* **2005,** *220*, 499.

(21) Létard, J.-F.; Guionneau, P.; Rabardel, L.; Howard, J. A. K.; Goeta, A. E.; Chasseau, D.; Kahn, O. Structural, Magnetic, and Photomagnetic Studies of a Mononuclear Iron(II) Derivative Exhibiting an Exceptionally Abrupt Spin Transition. Light-Induced Thermal Hysteresis Phenomenon. *Inorganic Chemistry* **1998,** *37* (17), 4432.

(22) Létard, J.-F.; Guionneau, P.; Codjovi, E.; Lavastre, O.; Bravic, G.; Chasseau, D.; Kahn, O. Wide Thermal Hysteresis for the Mononuclear Spin-Crossover Compound cis-Bis(thiocyanato)bis[N-(2'-pyridylmethylene)-4-(phenylethynyl)anilino]iron(II). *Journal of the American Chemical Society* **1997,** *119* (44), 10861.

(23) Perez-Mato, J.; Orobengoa, D.; Tasci, E.; De la Flor Martin, G.; Kirov, A. Crystallography Online: Bilbao Crystallographic Server. *Bulgarian Chemical Communications* **2011,** *43*, 183.

(24) Aroyo, M. I.; Perez-Mato, J. M.; Capillas, C.; Kroumova, E.; Ivantchev, S.; Madariaga, G.; Kirov, A.; Wondratschek, H. Bilbao Crystallographic Server: I. Databases and crystallographic computing programs. **2006,** *221* (1), 15.

(25) Aroyo, M. I.; Kirov, A.; Capillas, C.; Perez-Mato, J. M.; Wondratschek, H. Bilbao Crystallographic Server. II. Representations of crystallographic point groups and space groups. *Acta Crystallographica Section A* **2006,** *62* (2), 115.

(26) Gavezzotti, A. *Molecular Aggregation - Structure Analysis and Molecular Simulation of Crystals and Liquids*; 1st ed.; Oxford University Press: New York, 2007.

(27) Gavezzotti, A. Efficient computer modeling of organic materials. The atom–atom, Coulomb–London–Pauli (AA-CLP) model for intermolecular electrostatic-polarization, dispersion and repulsion energies. *New J. Chem.* **2011,** *35* (7), 1360.

(28) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.et al. Wallingford, CT, 2016.

(29) Reeves, M. G.; Wood, P. A.; Parsons, S. MrPIXEL: automated execution of Pixel calculations via the Mercury interface. *Journal of Applied Crystallography* **2020,** *53* (4), 1154.

(30) Gavezzotti, A. In *Zeitschrift für Kristallographie - Crystalline Materials*, 2005; Vol. 220.

(31) Maloney, A. G. P.; Wood, P. A.; Parsons, S. Intermolecular interaction energies in transition metal coordination compounds. *CrystEngComm* **2015,** *17* (48), 9300.

(32) Macrae, C. F.; Sovago, I.; Cottrell, S. J.; Galek, P. T. A.; McCabe, P.; Pidcock, E.; Platings, M.; Shields, G. P.; Stevens, J. S.; Towler, M.et al. Mercury 4.0: from

visualization to analysis, design and prediction. *Journal of Applied Crystallography* **2020,** *53* (1), 226.

(33) Mackenzie, C. F.; Spackman, P. R.; Jayatilaka, D.; Spackman, M. A. CrystalExplorer model energies and energy frameworks: extension to metal coordination compounds, organic salts, solvates and open-shell systems. *IUCrJ* **2017,** *4* (5), 575.

(34) Turner, M. J.; Thomas, S. P.; Shi, M. W.; Jayatilaka, D.; Spackman, M. A. Energy frameworks: insights into interaction anisotropy and the mechanical properties of molecular crystals. *Chemical Communications* **2015,** *51* (18), 3735.

(35) Bryant, M. J.; Maloney, A. G. P.; Sykes, R. A. Predicting mechanical properties of crystalline materials through topological analysis. *CrystEngComm* **2018,** *20* (19), 2698.

(36) Senthil Kumar, K.; Ruben, M. Emerging trends in spin crossover (SCO) based functional materials and devices. *Coordination Chemistry Reviews* **2017,** *346*, 176.

(37) Gajda, R.; Domański, M. A.; Malinska, M.; Makal, A. Crystal morphology fixed by interplay of π-stacking and hydrogen bonds – the case of 1-hydroxypyrene. *CrystEngComm* **2019,** *21* (11), 1701.

5. Concluding Remarks

The use of computational methods in interpreting intermolecular interactions within crystal structures has garnered increasing interest over the last decade. This is particularly true for small organic systems, where extensive work has already been done in rationalizing the crystal packing of structures using software such as PIXEL-CLP and CrystalExplorer. The use of these techniques when discussing the structures of coordination compounds has also grown all-be-it more slowly, in-part due to the increased computational resources needed, and the complexity based entry barrier to setting up calculations. With the introduction of the CSD Python API, it is now possible to automate the setup and execution of many of these calculations and demonstrate some of the invaluable insight the results can give towards the properties of inorganic crystal structures.

The first task, in implementing an automated BVS methodology within the CSD program Mercury, has resulted in a process that automates oxidation state assignment to structures in the CSD, as well as user-structures. This process has validated the reliability of BVS, showing that the correct oxidation state can be assigned using this method for 90% of entries, while also introducing new methods for increased confidence. The combination of techniques resulted in a process that can automatically assign oxidation states for over 80% of entries due to a near perfect accuracy when all methods agree. Ultimately, this has helped the Cambridge Structural Data Centre to retrospectively add over 50,000 oxidation state values to CSD entries over the last few years and aided them in ensuring this data is included in newly deposited structures. Additionally, it opens up the capability to add oxidation states as atom specific information, further improving the user experience of the CSD.

From here it has been possible to automate other computational techniques, with a specific aim to produce scripts that execute PIXEL-CLP from within Mercury. As previously alluded to, PIXEL calculations have been used extensively for organic structures, and with the recent parameterisation of transition metal atoms, it is now possible to extend its use to inorganic systems. The development of MrPixel, a software package designed to both set-up and run the PIXEL procedure on structures

from the Mercury interface, has enabled users to quickly gain more information on the intermolecular interactions within their crystal-structures. MrPixel has simplified the manually intensive process of setting up Pixel calculations, enabling it to become a routine process that can be employed on multiple structures simultaneously. This has already received significant interest, with many users already utilizing MrPixel and thus the PIXEL-CLP software in their structural analysis.

It is this development that has allowed PIXEL to be practicably employed on spin-crossover complexes. The various magnetic and electronic characteristics of SCO structures have long been a topic of debate, with differing behaviours between seemingly similar inorganic complexes being difficult to rationalize from the geometric data alone. By introducing computational techniques such as PIXEL-CLP, intermolecular energies can now be used to discuss and compare the differences between structures and spin-states, revealing trends between intermolecular interactions and the abruptness of spin-transition, and providing further evidence and understanding of the commonly discussed "co-operativity pathways". This work opens a significant avenue of further investigation for SCO chemistry and may be of significant benefit to those interested in designing new materials with desired properties.

This work demonstrates that determination of intermolecular energies for inorganic structures can provide a wealth of chemical and physical information. Each process has exemplified the use of the CSD Python API, which has proved invaluable in automating processes for crystal structure analysis. It is hoped that the results displayed further encourage users to explore the capabilities of both PIXEL and the CSD Python API, with the potential to add to and complement the vast functionality of the CSD package.

The Interpretation of the Crystal Structures of Transition Metal Compounds and Complexes

Thesis Supplementary Information

Matthew G Reeves

Contents

SI.2 Automated oxidation-state assignment for metal sites in coordination complexes in the Cambridge Structural Database.

Table S2.1: Thresholds *for common oxidation states. True values are those where >15% entries in the CSD have that named value. Bold are common while blue are those observed as defined by C. Housecroft & A. Sharpe* [1].

15% Threshold		Valence							
		i	ii	iii	iv	v	vi	vii	Viii
Metal	*Sc*	FALSE	FALSE	**TRUE**	FALSE	FALSE	FALSE	FALSE	FALSE
	Ti	FALSE	FALSE	FALSE	**TRUE**	FALSE	FALSE	FALSE	FALSE
	V	FALSE	FALSE	TRUE	**TRUE**	**TRUE**	**FALSE**	FALSE	FALSE
	Cr	FALSE	**TRUE**	**TRUE**	FALSE	FALSE	**FALSE**	FALSE	FALSE
	Mn	FALSE	**TRUE**	TRUE	FALSE	FALSE	FALSE	**FALSE**	FALSE
	Fe	FALSE	**TRUE**	**TRUE**	FALSE	FALSE	FALSE	FALSE	FALSE
	Co	FALSE	**TRUE**	**TRUE**	FALSE	FALSE	FALSE	FALSE	FALSE
	Ni	FALSE	**TRUE**	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
	Cu	**TRUE**	**TRUE**	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
	Zn	FALSE	**TRUE**	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
	Y	FALSE	FALSE	**TRUE**	FALSE	FALSE	FALSE	FALSE	FALSE
	Zr	FALSE	FALSE	FALSE	**TRUE**	FALSE	FALSE	FALSE	FALSE
	Nb	FALSE	FALSE	FALSE	TRUE	**TRUE**	FALSE	FALSE	FALSE
	Mo	FALSE	FALSE	FALSE	TRUE	FALSE	**TRUE**	FALSE	FALSE
	Tc	FALSE	FALSE	TRUE	**FALSE**	**TRUE**	FALSE	**FALSE**	FALSE
	Ru	FALSE	TRUE	**FALSE**	FALSE	FALSE	FALSE	FALSE	FALSE
	Rh	TRUE	FALSE	**TRUE**	FALSE	FALSE	FALSE	FALSE	FALSE
	Pd	FALSE	**TRUE**	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
	Ag	**TRUE**	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
	Cd	FALSE	**TRUE**	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
	Hf	FALSE	FALSE	FALSE	**TRUE**	FALSE	FALSE	FALSE	FALSE
	Ta	FALSE	FALSE	FALSE	FALSE	**TRUE**	FALSE	FALSE	FALSE
	W	FALSE	TRUE	FALSE	**TRUE**	**FALSE**	**TRUE**	FALSE	FALSE
	Re	TRUE	FALSE	**FALSE**	**FALSE**	**TRUE**	**FALSE**	**FALSE**	FALSE
	Os	FALSE	TRUE	FALSE	**TRUE**	FALSE	TRUE	FALSE	FALSE
	Ir	TRUE	FALSE	**TRUE**	**FALSE**	FALSE	FALSE	FALSE	FALSE
	Pt	FALSE	**TRUE**	FALSE	**FALSE**	FALSE	FALSE	FALSE	FALSE
	Au	TRUE	FALSE	**TRUE**	FALSE	FALSE	FALSE	FALSE	FALSE
	Hg	**FALSE**	**TRUE**	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE

For a full description of how BVS applied to all possible oxidation states, the process is demonstrated here for the four-co-ordinate cobalt complex KUYHES [1]. Application of BVS to each valence is described with a separate table showing how multiple bond valence sums are produced from each of the parameter values available. Each table shows top to bottom combinations which result in seperate bond valence sum values. For each bond encountered, all possible parameters are applied, producing multiple bond valence results. As we iterate through each sucessive bond, a new set of possible pond valences is produced and so a new possible bond valence sum value is also determined. The "best" bond valence sum value, along with the parameter "path" taken to arrive at it are highlighted in green. Note the single set of parameters available for Co(III) in this case, and therefore the single route and BVS value available in Table S3.

Table S2.2: Possible bond valence combinations for Co(II) assumed valence in CSD Refcode KUYHES.

Co-Cl1 r	2.219															
R0 Parameter	2.033								2.010							
Result	0.605								0.568							
Co-Cl2 r	2.227															
R0 Parameter	2.033				2.010				2.033				2.010			
Result	0.592				0.556				0.592				0.556			
Co-N1 r	2.042															
R0 Parameter	1.720		1.600		1.720		1.600		1.720		1.600		1.720		1.600	
Results	0.419		0.303		0.419		0.303		0.419		0.303		0.419		0.303	
Co-N2 r	2.053															
R0 Parameter	1.720	1.600	1.720	1.600	1.720	1.600	1.720	1.600	1.720	1.600	1.720	1.600	1.720	1.600	1.720	1.600
Results	0.407	0.294	0.407	0.294	0.407	0.294	0.407	0.294	0.407	0.294	0.407	0.294	0.407	0.294	0.407	0.294
Bond Valence Sum	2.022	1.910	1.906	1.794	1.987	1.874	1.871	1.758	1.986	1.873	1.870	1.757	1.950	1.838	1.834	1.721
Delta	0.022	0.090	0.094	0.206	0.013	0.126	0.129	0.242	0.014	0.127	0.130	0.243	0.050	0.162	0.166	0.279

Table S2.3: *Possible bond valence combination for Co(III) assumed valence in CSD Refcode KUYHES.*

Co-Cl1 r	2.219
R0 Parameter	2.050
Result	0.633
Co-Cl2 r	2.227
R0 Parameter	2.050
Result	0.620
Co-N1 r	2.042
R0 Parameter	1.690
Results	0.386
Co-N2 r	2.053
R0 Parameter	1.690
Results	0.375
Bond Valence Sum	2.014
Delta	0.986

Table S2.4: Breakdown of confidence scores for oxidation state assignment in the test dataset.

Confidence	Correct	Incorrect	Total	% Correct
0	0	3113	3113	0.00%
4	8	56	64	12.50%
5	533	2399	2932	18.18%
6	1184	1781	2965	39.93%
7	425	124	549	77.41%
8	891	63	954	93.40%
9	1130	34	1164	97.08%
10	2740	85	2825	96.99%
11	3407	55	3462	98.41%
12	873	18	891	97.98%
13	1106	6	1112	99.46%
14	2363	19	2382	99.20%
15	6203	35	6238	99.44%
16	13038	56	13094	99.57%
17	13171	83	13254	99.37%
Total	47072	7927	54999	85.59%
Summary:				
Band	**Correct**	**Incorrect**	**Total**	**% Correct**
A	35881	199	36080	99.45%
B	8150	192	8342	97.70%
C	2500	1968	4468	55.95%
D	541	2455	2996	18.06%

SI.3 MrPIXEL: Execution of PIXEL Calculations via the MERCURY Interface

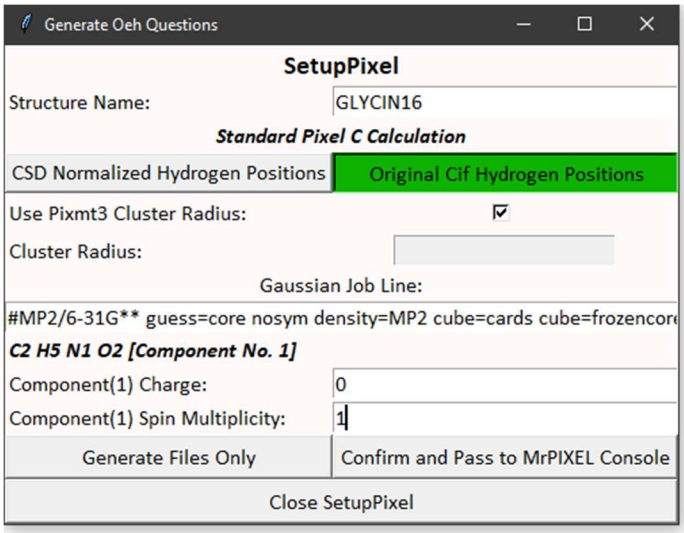
Table S3.1: Pixel identifiers for different atom types. Mol2 format SYBYL notations are included, as well as some custom labels denoted by * assigned by SetupPixel. This table is based on that in the Pixel manual.

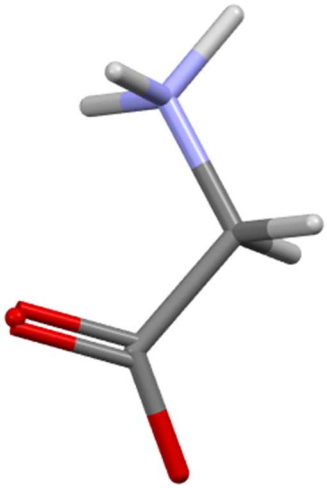
Atom descriptor	***.mol2* SYBYL notation**	***.oeh* indicator**	**Atomic Polarizability (Å^3)**	**Ionization Potential (a.u.)**	**Space diffusion Parameter**	**H-bond Propensity**
hydrogen radius 1.10			0.39	0.5		
acetylene CH	**H.1**	**1**			**0.6**	**0.2**
=CH_2, arom.CH	**H.2**	**2**			**0.62**	**0.1**
aliphatic CH, CH_2, CH_3	**H.3**	**3**			**0.64**	**0.05**
R-OH, R-SH alcohol, thiol	**H.5**	**5**			**0.75**	**0.99**
COO-H acid	**H.6**	**6**			**0.8**	**0.99**
CON)-H amide	**H.am**	**7**			**0.8**	**0.9**
R_2NH, RNH_2, (R_3N+)H	**H.8**	**8**			**0.8**	**0.99**
H2O (water)	**H2.o**	**9**			**0.8**	**0.99**
unnormalized hydrogen atom from Cambridge files		**99**				
carbon 1.77				0.414	1	0
carbonyl C=(O)	**C.O**	**10**	**1.05**			
≡C-	**C.1**	**11**	**1.35**			
sp^2 or allene C	**C.2, C.ar**	**12**	**1.35**			
sp^3 C	**C.3**	**13**	**1.05**			
aromatic core C	**C.Core**	**14**	**1.9**			
nitrogen 1.64			0.95	0.534		
(RnH_4-n)N+	**N.4**	**16**			**0.63**	**0**
(RnH_3-n)N	**N.3**	**17**			**0.63**	**-0.97**
arom.N, R=N(H)	**N.ar,N.pl3**	**18**			**0.58**	**-0.99**
-C≡N,-N=N	**N.2**	**19**			**0.7**	**-0.7**

Atom descriptor		***.mol2* SYBYL notation**	***.oeh* indicator**	**Atomic Polarizability (Å^3)**	**Ionization Potential (a.u.)**	**Space diffusion Parameter**	**H-bond Propensity**
nitro N		**N.1**	**20**			**0.63**	**0**
amide N (CONH,$CONH_2$)		**N.am**	**21**			**0.63**	**-0.85**
oxygen	**1.58**			0.75	0.5		
-O-		**O.3**	**23**			**0.45**	**-0.9**
H_2O (water)		**O.1**	**24**			**0.7**	**-0.99**
C=O, COO-		**O.co2, O.4,O.2**	**27**			**0.5**	**-0.99**
(C=O)-OH		**O.car**	**28**			**0.5**	**-0.9**
R-OH		**O.al**	**29**			**0.45**	**-0.99**
N=O		**O.2**	**30**			**0.5**	**-0.95**
S=O		**O.2**	**31**			**0.75**	**-0.9**
P=O		**O.3**	**32**			**0.75**	**-0.9**
sulfur	**1.81**			3	0.381		
-S-		**S.3**	**34**			**2**	**-0.5**
(C)=S		**S.1**	**35**			**2**	**-0.5**
(O)=S		**S, S.o2**	**36**			**2.5**	**0**
R-S(H)		**S.2**	**37**			**2**	**-0.5**
heteroatoms							
P	1.9	**P**	**38**	**1.54**	**0.386**	**3**	**0**
As	1.8	**As**	**39**	**3.5**	**0.4**	**5**	**0**
Se	1.8	**Se**	**40**	**3.5**	**0.4**	**6**	**0**
F	1.46	**F.0**	**41**	**0.55**	**0.64**	**0.2**	**0**
Cl	1.76	**Cl.0**	**42**	**2.5**	**0.477**	**2.4**	**-0.2**
Br	1.87	**Br.0**	**43**	**3.27**	**0.434**	**1.5**	**0**
I	2.03	**I.0**	**44**	**5**	**0.384**	**5**	**0**
B	1.70	**B**	**46**	**1.6**	**0.305**	**1**	**0**
Si	2.00	**Si**	**47**	**2**	**0.28**	**1**	**0**
transition metals							
Ti		**Ti**	**51**	**4.18**	**0.25**	**0.8**	**-0.5**

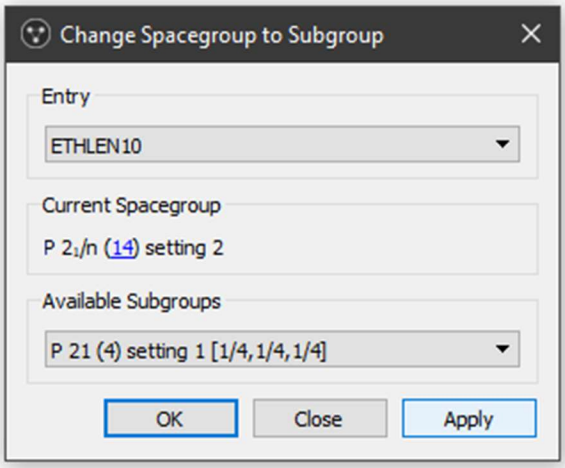
Atom descriptor	*.mol2* SYBYL notation	*.oeh* indicator	Atomic Polarizability (Å^3)	Ionization Potential (a.u.)	Space diffusion Parameter	H-bond Propensity
V	**V**	**52**	**3.31**	**0.25**		
Cr	**Cr, Cr.oh, Cr.th**	**53**	**2.86**	**0.25**		
Mn	**Mn**	**54**	**2.93**	**0.27**		
Fe	**Fe**	**55**	**2.81**	**0.29**		
Co	**Co, Co.oh**	**56**	**2.62**	**0.29**		
Ni	**Ni**	**57**	**2.61**	**0.28**		
Cu	**Cu**	**58**	**2.81**	**0.285**		
Zn	**Zn**	**59**	**3.63**	**0.345**		
positive ions						
Li+	**Li**	**61**	**0.1**	**1**	**0.2**	**0**
Na+	**Na**	**62**	**0.2**	**0.85**	**0.3**	
K+	**K**	**63**	**0.3**	**0.7**	**1.5**	
Rb+	**Rb**	**64**	**0.4**	**0.5**	**3**	
Cs+	**Cs**	**65**	**0.3**	**0.45**	**5**	
Ca+	**Ca**	**66**	**0.7**	**0.7**	**1.5**	
negative ions						
F-	**F.-I**	**67**	**0.4**	**0.75**	**0.5**	
Cl-	**Cl.-I**	**68**	**2.5**	**0.65**	**3**	
Br-	**Br.-I**	**69**	**3.27**	**0.5**	**4**	
I-	**I.-I**	**70**	**5**	**0.4**	**5**	

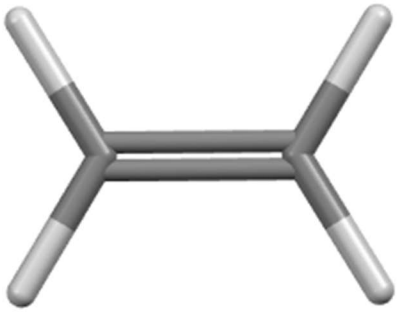
i.


Generate Oeh Questions
SetupPixel
Structure Name:
GLYCIN16
Standard Pixel C Calculation
CSD Normalized Hydrogen Positions
Original Cif Hydrogen Positions
Use Pixmt3 Cluster Radius:
Cluster Radius:
Gaussian Job Line:
#MP2/6-31G** guess=core nosym density=MP2 cube=cards cube=frozencor
C2 H5 N1 O2 [Component No. 1]
Component(1) Charge:
0
Component(1) Spin Multiplicity:
1
Generate Files Only
Confirm and Pass to MrPIXEL Console
Close SetupPixel

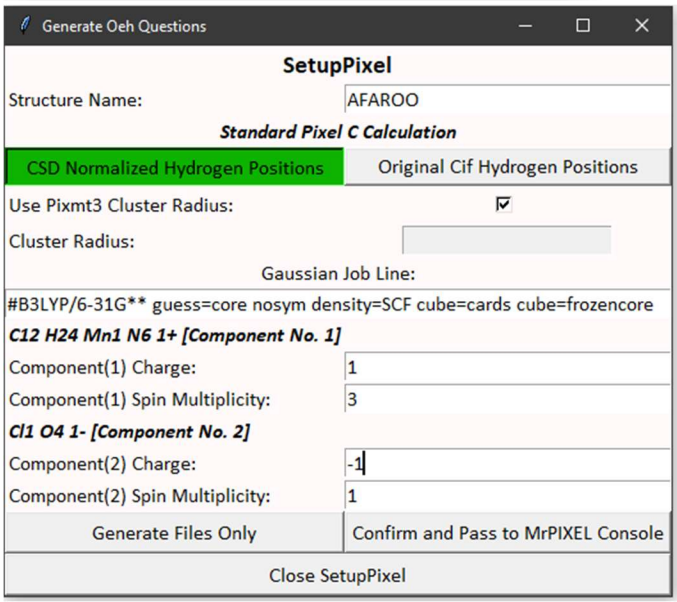


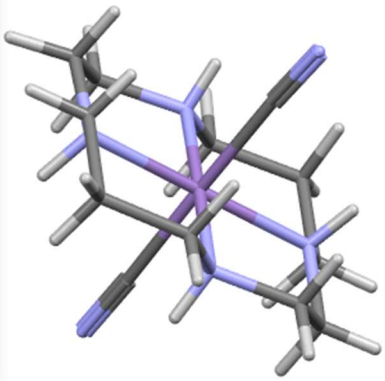
ii.


Change Spacegroup to Subgroup
Entry
ETHLEN10
Current Spacegroup
P 2₁/n (14) setting 2
Available Subgroups
P 21 (4) setting 1 [1/4,1/4,1/4]
OK
Close
Apply



iii.


Generate Oeh Questions
SetupPixel
Structure Name:
AFAROO
Standard Pixel C Calculation
CSD Normalized Hydrogen Positions
Original Cif Hydrogen Positions
Use Pixmt3 Cluster Radius:
Cluster Radius:
Gaussian Job Line:
#B3LYP/6-31G** guess=core nosym density=SCF cube=cards cube=frozencore
C12 H24 Mn1 N6 1+ [Component No. 1]
Component(1) Charge:
1
Component(1) Spin Multiplicity:
3
Cl1 O4 1- [Component No. 2]
Component(2) Charge:
-1
Component(2) Spin Multiplicity:
1
Generate Files Only
Confirm and Pass to MrPIXEL Console
Close SetupPixel



iv.

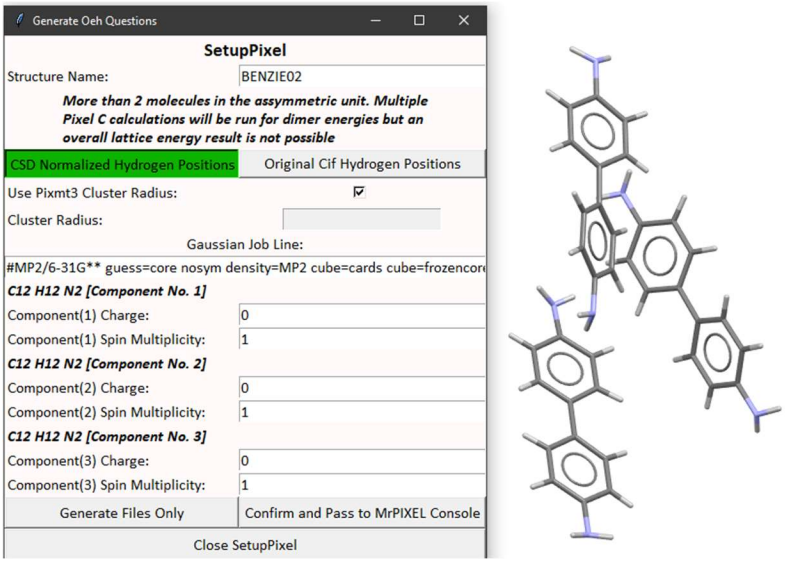


Figure S3.1: Example menus of SetupPixel and MRPixel. (i). SetupPIXEL for GLYCIN16 structure, (ii) symmetry lowering for ethylene, (iii) a two-component system where the cation is a transition metal complex, and (iv) a system for which Z' = 3.

Table S3.2: Breakdown of Pixel-C results for first coordination sphere of γ-glycine (GLYCIN16 CSD structure). Note small discrepancy (0.1 kJ mol^{-1} difference) in the pair of interactions at a distance of 4.389 Å is due to rounding errors.

Symmetry operation	Centroid-centroid distance (Å)	E_{Elec} (kJ mol^{-1})	E_{POL} (kJ mol^{-1})	E_{DISP} (kJ mol^{-1})	E_{REP} (kJ mol^{-1})	E_{TOT} (kJ mol^{-1})
x,*y*,*z*-1	5.473	−119.2	−38.6	−14.3	66.3	−105.8
x,*y*,*z*+1	5.473	−119.2	−38.6	−14.3	66.3	−105.8
-*x*+*y*,-*x*-1,*z*+⅔	5.458	−37.5	−11.2	−8.1	9.5	−47.5
-*y*+1,*x*-*y*,*z*-⅔	5.458	−37.5	−11.2	−8.1	9.5	−47.5
-*x*+*y*,-*x*-1,*z*-⅓	4.450	49.7	−8.7	−10.5	7.1	37.7
-*y*+1,*x*-*y*,*z*+⅓	4.450	49.7	−8.7	−10.5	7.1	37.7
-*y*+1,*x*-*y*,*z*-⅓	4.389	−28.3	−41.0	−18.5	54.4	−33.4
-*x*+*y*+1,-*x*+1,*z*+⅓	4.389	−28.3	−41.0	−18.5	54.4	−33.3
-*y*,*x*-*y*,*z*+⅔	5.437	−23.6	−11.0	−7.2	10.2	−31.6
-*y*,*x*-*y*,*z*-⅔	5.437	−23.6	−11.0	−7.2	10.2	−31.6
-*x*+*y*,-*x*,*z*+⅓	4.424	29.7	−33.7	−17.1	33.3	12.2
-*y*,*x*-*y*,*z*-⅓	4.424	29.7	−33.7	−17.1	33.3	12.2
-*y*+1,*x*-*y*,*z*+⅔	5.408	11.0	−4.4	−3.6	1.0	4.0
-*y*+1,*x*-*y*+1,*z*-⅔	5.408	11.0	−4.4	−3.6	1.0	4.0

Table S3.3: Breakdown of Pixel results for first coordination sphere of ethylene (ETHLEN10 CSD structure).

Symmetry operation	Centroid-centroid distance (Å)	E_{Elec} (kJ mol^{-1})	E_{POL} (kJ mol^{-1})	E_{DISP} (kJ mol^{-1})	E_{REP} (kJ mol^{-1})	E_{TOT} (kJ mol^{-1})
x,*y*, *z*−1	4.067	-1.4	-0.4	-5.8	3.0	-4.7
x,*y*, *z*+1	4.067	-1.4	-0.4	-5.8	3.0	-4.7
−*x*+½,*y*−½,−*z*+½	4.441	-0.4	-0.2	-4.7	2.1	-3.2
−*x*+½,*y*+½,−*z*+½	4.441	-0.4	-0.2	-4.7	2.1	-3.2
−*x*−½,*y*−½,−*z*−½	4.441	-0.4	-0.2	-4.8	2.1	-3.3
−*x*−½,*y*+½,−*z*−½	4.441	-0.4	-0.2	-4.8	2.1	-3.3
−*x*−½,*y*−½,−*z*+½	4.600	-1.4	-0.3	-2.9	1.7	-3.0
−*x*−½,*y*+½,−*z*+½	4.600	-1.4	-0.3	-2.9	1.7	-3.0
−*x*+½,*y*−½,−*z*−½	4.600	-1.5	-0.3	-2.9	1.7	-3.1
−*x*+½,*y*+½,−*z*−½	4.600	-1.5	-0.3	-2.9	1.7	-3.1
x−1,*y*, *z*	4.626	-0.7	-0.2	-3.5	1.8	-2.6
x+1,*y*, *z*	4.626	-0.7	-0.2	-3.5	1.8	-2.6

Table S3.4: The first coordination sphere of the cations in AFAROO. M = central reference cation, M1, M2… are cations related by the operations listed. A1, A2 etc are anions.

Contact (see Fig. 4)	Symmetry operation	Centroid-Centroid distance (Å)	E_{Elec} (kJ mol^{-1})	E_{POL} (kJ mol^{-1})	E_{DISP} (kJ mol^{-1})	E_{REP} (kJ mol^{-1})	E_{TOT} (kJ mol^{-1})
M...M1	$x-1,y,z$	6.760	−60.6	−33.2	−36.7	64.0	−66.5
M...M2	$x+1,y,z$	6.760	−60.6	−33.2	−36.7	64.0	−66.5
M...M3	$-x, y-½,-z+3/2$	8.668	−8.1	−4.8	−15.0	8.8	−19.1
M...M4	$-x,y+½,-z+3/2$	8.668	−8.1	−4.8	−15.0	8.8	−19.1
M...M5	$x-½,-y+½,-z+2$	8.502	0.2	−4.2	−13.8	5.4	−12.4
M...M6	$x+½,-y+½,-z+2$	8.502	0.2	−4.2	−13.8	5.4	−12.4
M...M7	$x-½,-y-½,-z+2$	9.442	−1.6	−1.1	−10.5	3.8	−9.4
M...M8	$x+½,-y-½,-z+2$	9.442	−1.6	−1.1	−10.5	3.8	−9.4
M...M9	$-x+1,y-½,-z+3/2$	8.900	−0.0	−2.3	−15.8	9.0	−9.1
M...M10	$-x+1,y+½,-z+3/2$	8.900	−0.0	−2.3	−15.8	9.0	−9.1
M...A1	x,y,z	5.359	−14.7	−8.2	−26.5	33.0	−16.3
M...A2	$-x+½,-y,z+½$	5.371	−11.5	−6.9	−27.1	29.7	−15.8
M...A3	$-x,y-½,-z+3/2$	6.546	−0.6	−1.8	−14.8	7.6	−9.7
M...A4	$-x,y+½,-z+3/2$	6.928	−0.0	−1.3	−13.2	6.5	−8.0
M...A5	$x+1,y,z$	6.632	−1.1	−0.8	−10.2	3.5	−8.5
M...A6	$-x-½,-y,z+½$	6.327	−2.7	−1.7	−14.9	9.4	−9.9

Table S3.5: Comparison of strongest interaction energies in the crystal structure of acetoxime. Energies are in kJ mol^{-1}.

Molecule a			***Molecule b***			***Molecule c***		
Dimer	**Distance (Å)**	E_{TOT}	**Dimer**	**Distance (Å)**	E_{TOT}	**Dimer**	**Distance (Å)**	E_{TOT}
b	5.935	−28.1	c	5.942	−27.6	a	5.964	−28.3
c	5.964	−28.3	a	5.935	−28.1	b	5.942	−27.6
b[*x*,*y*−1,*z*]	6.077	−4.4	c[*x*,*y*+1,*z*−1]	6.187	−4.1	a[*x*,*y*,*z*+1]	6.102	−4.3
c[*x*,*y*,*z*−1]	6.102	−4.3	a[*x*,*y*+1,*z*]	6.077	−4.4	b[*x*,*y*−1,*z*+1]	6.187	−4.1
b[*x*,*y*,*z*−1]	6.274	−3.7	c[*x*,*y*+1,*z*]	6.170	−4.5	a[*x*,*y*−1,*z*+1]	6.238	−4.0
c[*x*,*y*+1,*z*−1]	6.238	−4.0	a[*x*,*y*,*z*+1]	6.274	−3.7	b[*x*,*y*−1,*z*]	6.170	−4.5
b[−*x*−1,−*y*,−*z*]	5.033	−7.4	c[−*x*−1,−*y*−1,−*z*+1]	4.631	−10.8	a[−*x*−1,−*y*−1,−*z*]	5.354	−6.7
a[−*x*−1,−*y*−1,−*z*]	4.649	−10.3	b[−*x*−1,−*y*,−*z*]	5.356	−7.5	c[−*x*−1,−*y*−1,−*z*+1]	4.540	−10.6
c[−*x*−1,−*y*−1,−*z*]	5.354	−6.7	a[−*x*−1,−*y*,−*z*]	5.033	−7.3	b[−*x*−1,−*y*−1,−*z*+1]	4.631	−10.8
b[−*x*,−*y*,−*z*]	5.394	−7.1	c[−*x*,−*y*−1,−*z*+1]	5.137	−7.4	a[−*x*,−*y*−1,−*z*]	4.565	−11.3
a[−*x*,−*y*−1,−*z*]	4.754	−9.2	b[−*x*,−*y*,−*z*]	4.299	−12.8	c[−*x*,−*y*−1,−*z*+1]	5.543	−5.9
c[−*x*,−*y*−1,−*z*]	4.565	−11.3	a[−*x*,−*y*,−*z*]	5.394	−7.1	b[−*x*,−*y*−1,−*z*+1]	5.137	−7.4

SI.4 Revealing Cooperativity Pathways in Spin Crossover Complexes

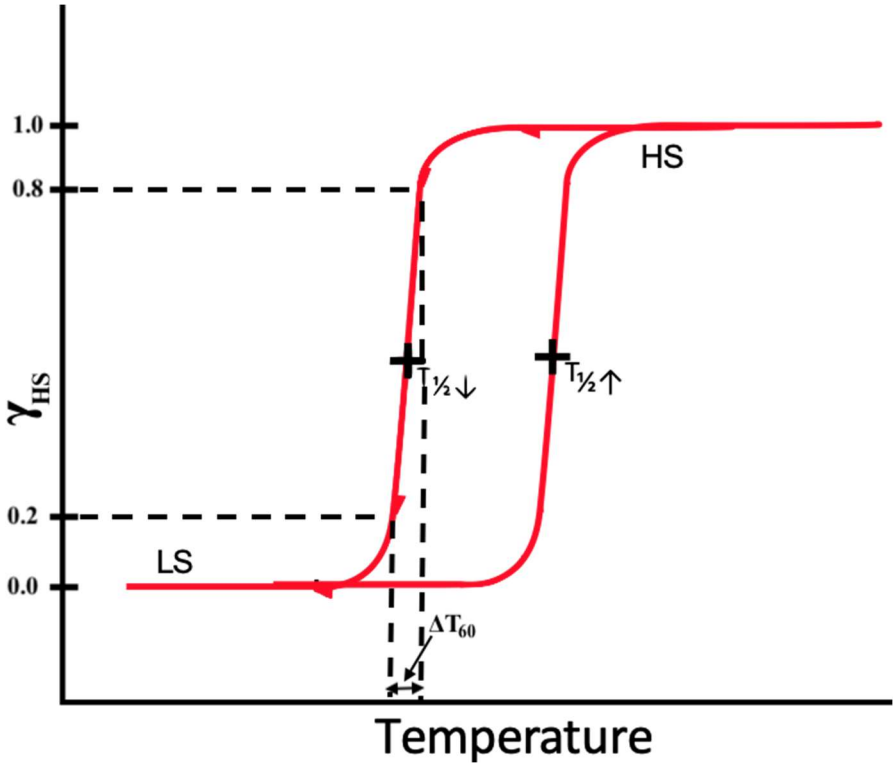


Figure S4.1: Aspects of temperature−induced SCO behaviour for a SCO material exhibiting a transition with hysteresis. Arrows show direction of spin transition. In case of hysteresis, two values of $T_{1/2}$ exist, corresponding to the ascending (↑) and descending (↓) branches. The hysteresis width is usually denoted $\Delta T_{1/2}$ (where $\Delta T_{1/2} = T_{1/2}\uparrow - T_{1/2}\downarrow$). The abruptness of the transition represents the temperature difference between the 20 and 80% HS state rate and is denoted ΔT_{60}. Where a hysteresis is present, two values of ΔT_{60} exist according to the sharpness of the transition during the cooling and the warming modes. Here only the HS−>LS transition abruptness $\Delta T_{60}\downarrow$ is shown.

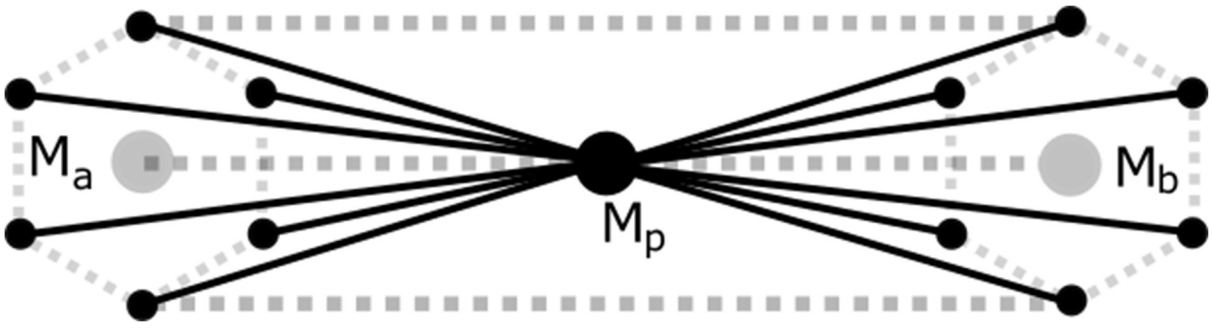


Figure S4.2: Example of strut construction showing new atoms placed between Molecule M_a and M_b and bonds between newly placed atoms. Magnitude of hexagons dictated by energy over value n where E/n produces a sensible strut size.

Packing in the crystal structures of Fe(PM-L)2(NCS)2 complexes

Shown here are high-spin $Fe(PM\text{-}NeA)_2(NCS)_2$ and $Fe(PM\text{-}BiA)_2(NCS)_2$ polymorph-II which correspond to the smallest and largest layer separation respectively. Quantitative values of layer spacing are displayed in Table S1.

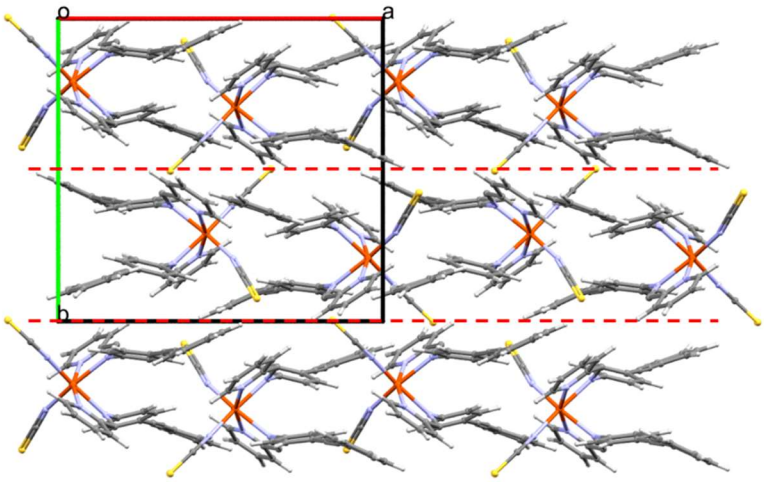


Figure S4.3: View of the layer structure of HS $Fe(PM\text{-}NeA)_2(NCS)_2$ along c.

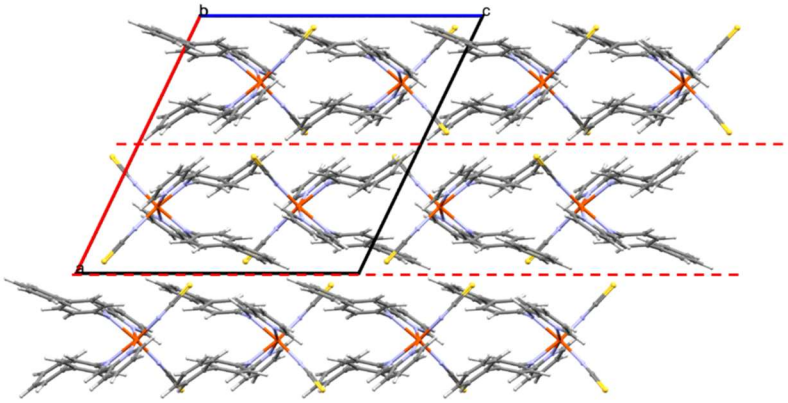


Figure S4.4: View of the layer structure of HS $Fe(PM\text{-}BiA)_2(NCS)_2$ polymorph-II along b.

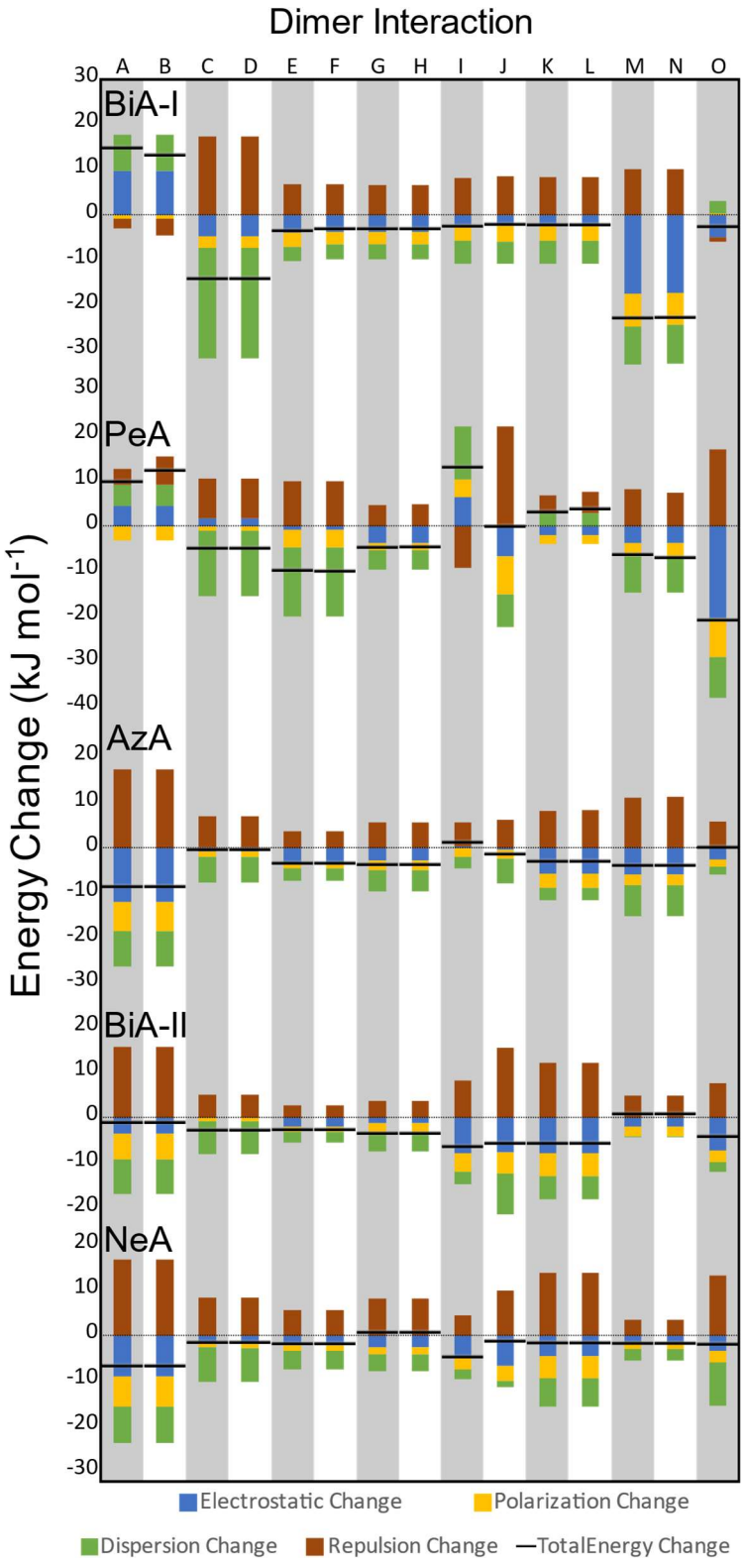
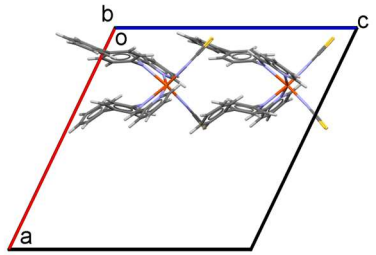
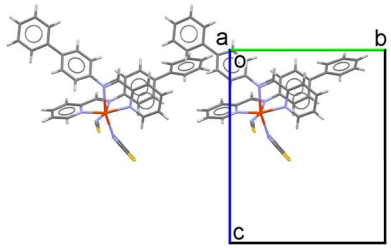
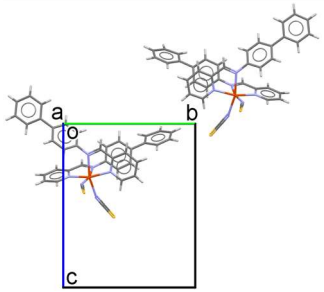


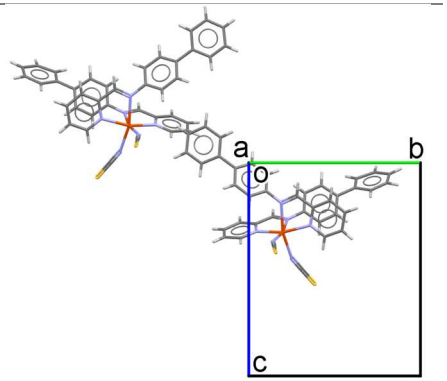
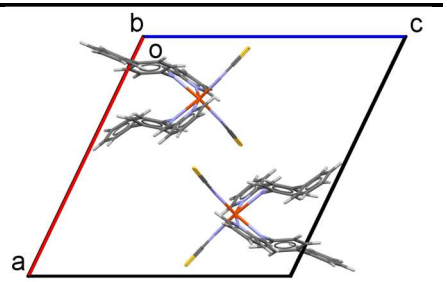
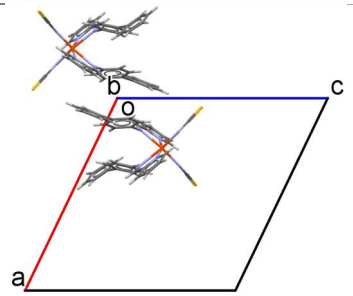
Figure S4.5: Stacked bar−graph of interaction energy changes from HS to LS structures for PM−L complexes, broken down by separate energy components. The total energy changes are shown as black bars: destabilising changes which appear as pink struts in the framework plots of Figs. 5 and 6 are located above the zero-energy; the stabilising changes represented as green struts in are below the axis.

Table S4.1: Layers in SCO structures. Values determined using the topological analysis scripts developed by Bryant et. al.[3]

Structure	Refcode	Slip Plane [h,k,l]	Layer Spacing (Å)
Fe(PM-AzA)$_2$(NCS)$_2$ HS	XECNAU35	[0,1,0]	0.10
Fe(PM-AzA)$_2$ (NCS)$_2$ LS	XECNAU07	[0,1,0]	0.02
Fe(PM-BiA)$_2$ (NCS)$_2$−I HS	RONPIT01	[0,2,0]	0.12
Fe(PM-BiA)$_2$ (NCS)$_2$−I LS	RONPIT02	[0,2,0]	1.40
Fe(PM-BiA)$_2$ (NCS)$_2$−II HS	RONPIT04	[1,0,0]	0.92
Fe(PM-BiA)$_2$ (NCS)$_2$−II LS	RONPIT05	[1,0,0]	1.06
Fe(PM-PeA)$_2$ (NCS)$_2$ HS	NOWBIK01	[0,1,0]	−0.10
Fe(PM-PeA)$_2$ (NCS)$_2$ LS	NOWBIK	[2,0,0]	0.50
Fe(PM-NeA)$_2$(NCS)$_2$ HS	COMQUR	[0,1,0]	−0.78
Fe(PM-(NeA)$_2$(NCS)$_2$ LS	COMQUR01	[0,1,0]	−0.74

Table S4.2: First co−ordination sphere interactions (in kJ mol^{-1}) for HS $Fe(PM\text{-}BiA)_2(NCS)_2$ polymorph-II.

Interaction	Centroid-Centroid Distance	Symmetry	E_{Elec}	E_{Pol}	E_{Disp}	E_{Rep}	E_{Tot}	Description	Figure
Intra−Layer									
A/B	8.719	x,−y+½,z−½ x,−y+½,z+½	−57.4	−28.8	−80.4	68.6	−97.9	Interaction along chain in a single layer with overlap of NCS and PM−X ligands	
C/D	12.602	x,y−1,z x,y+1,z	9.6	−5.9	−44.2	19.6	−20.9	Adjacent complexes in same layer with same orientation of ligands and molecular axis.	
E/F	15.996	x,3/2−y, z−1/2 x,3/2−y, z+1/2	−6.5	−5.0	−11.5	8.4	−14.6	Diagonal complexes in same layer with same molecular axis and opposite ligand orientations.	

G/H	14.621	x,−y−1/2,z+1/2 x,−y−1/2,z−1/2	−9.1	−7.2	−18.8	12.6	−22.6	Diagonal complexes in same layer with same molecular axis and opposite ligand orientations. Typically, shorter than E/F contacts.	
Inter−Layer									
I	12.318	1−x, −y,1−z	−39.0	−8.0	−12.6	7.9	−51.7	Diagonal contact between layers with thiocyanate ligands pointing away from the contact vector. 2 short NCS…PM contacts (4.035 Å).	
J	9.833	−x, −y, −z	−27.7	−14.7	−77.7	39.4	−80.7	Adjacent interlayer contact with large overlap of PM−R group aromatic rings, resulting in a strong dispersion term.	

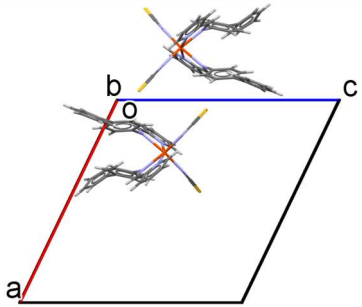
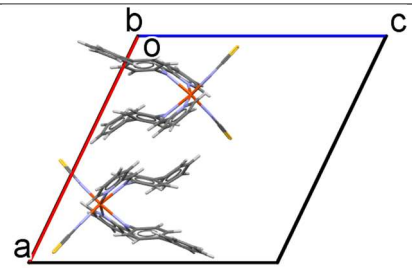
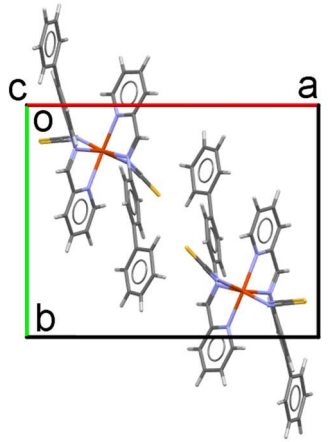
K/L	12.090	−x, y+1/2, 1/2−z −x, y−1/2, 1/2−z	−34.2	−12	−21.3	20.1	−47.5	Diagonal contact between layers with short NCS…R group contacts.	
M/N	10.171	1−x, y−1/2, 1/2−z 1−x, y+1/2, 1/2−z	−17.2	−11.1	−44.6	28.2	−44.7	Adjacent interlayer contact similar to Interaction J but with poorer aromatic ring stacking.	
O	13.145	1−x, 1−y, 1−z	−14.8	−12.6	−23.5	23.5	−27.4	Diagonal contact across layers with NCS groups orientated along the contact vector. Shortest C...S interaction.	

Table S4.3: Shortest C…S contact distances and energies in the HS forms.

PM−X Group	ΔT_{60} (K)	Shortest HS C…S Distance (Å)	Shortest LS C…S Distance (Å)	HS Symmetry Operation	HS Interaction Energy (kJ mol^{-1})	LS Interaction Energy (kJ mol^{-1})	ΔE_{LS-HS} (kJ mol^{-1})
BiA−I	5	3.417	3.436	2−x,−y,−z	−37.8	−40.6	−2.8
PeA	33.5	3.447	3.585	2−x,1−y,z	−21.0	−42.0	−21.0
AzA	60	3.484	3.411	1−x,−y,−z	−20.0	−20.1	−0.1
BiA−II	81	3.541	3.450	1−x,1−y,1−z	−27.4	−31.8	−4.4
NeA	97	3.438	3.445	−x,1−y,1−z	−31.6	−33.7	−2.1

Table S4.4: Overall PIXEL lattice energies (in kJ mol^{-1}) for the Fe(PM-L)$_2$(NCS)$_2$ family of SCO complexes.

PM−X Group	ΔT_{60} (K)	PIXEL Lattice Energies (40 Å Cut−off)		
		HS $E_{Lattice}$	LS $E_{Lattice}$	$\Delta E_{Lattice}$
BiA−I	5	−300.5	−327.8	−27.3
PeA	33.5	−300.4	−351.9	−51.5
AzA	60	−319.3	−342.6	−23.3
BiA−II	81	−297.9	−327.8	−29.9
NeA	97	−352.8	−372.7	−19.9

Table S4.5: Interaction−energy changes (in kJ mol^{-1}) in the HS → LS transition of $Fe(PM\text{-}L)_2(NCS)_2$ structures. The figures in the top and bottom rows are plotted in Fig.7 in the main text.

Parameter	Structure				
	BiA-I	**PeA**	**AzA**	**BiA-II**	**NeA**
ΔT_{60} **(K)**	5.0	33.5	60.0	81.0	97.0
Interaction Energy Changes ΔE_{TOT} **(kJ** mol^{-1}**)**					
A	14.6	9.6	−8.8	−1.3	−6.9
B	13.0	12.1	−8.8	−1.3	−6.9
C	−14.3	−5.1	−0.6	−3.0	−1.7
D	−14.3	−5.1	−0.6	−3.0	−1.7
E	−3.7	−10.0	−3.6	−2.9	−2.0
F	−3.3	−10.2	−3.6	−2.9	−2.0
G	−3.3	−4.9	−3.9	−3.7	0.5
H	−3.3	−4.8	−3.9	−3.7	0.5
I	−2.7	12.8	1.0	−6.6	−4.9
J	−2.3	−0.3	−1.6	−5.9	−1.4
K	−2.4	2.9	−3.2	−5.9	−1.8
L	−2.4	3.6	−3.2	−5.9	−1.8
M	−23.0	−6.5	−4.1	0.6	−1.9
N	−22.9	−7.2	−4.1	0.6	−1.9
O	−2.8	−21.0	−0.1	−4.4	−2.1
$\sum\Delta\lvert E_{Tot}\rvert$	128.3	116.1	51.1	51.7	38.0

Table S4.6: PIXEL-C interaction energy changes (HS>LS) for first intermolecular co-ordination sphere contacts of $Fe(PM\text{-}L)_2(NCS)_2$ complexes and other structures studied.

BiA-I	State	Interaction Transformation Matrix and Vector												Centroid Distances (Å)	E_{Elec}	E_{Pol}	E_{Disp}	E_{Rep}	E_{Tot}
															kJ mol^{-1}				
A	HS	1	0	0	0	−1	0	0	0	1	0	0.5	−0.5	8.804	−58.8	−28.1	−80.5	59.2	−108.1
	LS	1	0	0	0	−1	0	0	0	1	0	0.5	−0.5	9.140	−49.0	−28.9	−72.6	57.0	−93.5
	LS−HS													0.336	9.8	−0.8	7.9	−2.2	14.6
B	HS	1	0	0	0	−1	0	0	0	1	0	0.5	0.5	8.805	−58.8	−28.1	−80.5	61.1	−106.2
	LS	1	0	0	0	−1	0	0	0	1	0	0.5	0.5	9.141	−49.0	−28.9	−72.6	57.3	−93.2
	LS−HS													0.336	9.8	−0.8	7.9	−3.8	13.0
C	HS	1	0	0	0	1	0	0	0	1	−1	0	0	12.949	14.2	−4.9	−33.4	10.3	−13.8
	LS	1	0	0	0	1	0	0	0	1	−1	0	0	12.370	9.4	−7.4	−57.9	27.7	−28.1
	LS−HS													−0.579	−4.8	−2.5	−24.5	17.4	−14.3
D	HS	1	0	0	0	1	0	0	0	1	1	0	0	12.949	14.2	−4.9	−33.4	10.3	−13.8
	LS	1	0	0	0	1	0	0	0	1	1	0	0	12.370	9.4	−7.4	−57.9	27.7	−28.1
														−0.579	−4.8	−2.5	−24.5	17.4	−14.3
E	HS	1	0	0	0	−1	0	0	0	1	−1	0.5	0.5	15.659	−7.4	−5.5	−12.3	7.9	−17.2
	LS	1	0	0	0	−1	0	0	0	1	−1	0.5	0.5	15.381	−11.3	−8.6	−15.6	14.6	−20.9
	LS−HS													−0.278	−3.9	−3.1	−3.3	6.7	−3.7
F	HS	1	0	0	0	−1	0	0	0	1	1	0.5	0.5	15.659	−7.4	−5.5	−12.3	7.9	−17.2
	LS	1	0	0	0	−1	0	0	0	1	1	0.5	0.5	15.381	−11.1	−8.4	−15.5	14.6	−20.5
	LS−HS													−0.278	−3.7	−2.9	−3.2	6.7	−3.3
G	HS	1	0	0	0	−1	0	0	0	1	−1	0.5	−0.5	15.659	−7.4	−5.5	−12.3	8.0	−17.2
	LS	1	0	0	0	−1	0	0	0	1	−1	0.5	−0.5	15.381	−11.1	−8.4	−15.5	14.5	−20.5
	LS−HS													−0.278	−3.7	−2.9	−3.2	6.5	−3.3
H	HS	1	0	0	0	−1	0	0	0	1	1	0.5	−0.5	15.659	−7.4	−5.5	−12.3	7.9	−17.2
	LS	1	0	0	0	−1	0	0	0	1	1	0.5	−0.5	15.381	−11.1	−8.4	−15.5	14.5	−20.5
														−0.278	−3.7	−2.9	−3.2	6.6	−3.3
I	HS	−1	0	0	0	1	0	0	0	−1	2	−0.5	0.5	10.000	−19.4	−11.2	−48.3	30.0	−48.9
	LS	−1	0	0	0	1	0	0	0	−1	2	−0.5	0.5	9.703	−21.6	−14.8	−53.3	38.2	−51.6
	LS−HS													−0.297	−2.2	−3.6	−5.0	8.2	−2.7

BiA-I	State	Interaction Transformation Matrix and Vector												Centroid Distances (Å)	E_{Elec}	E_{Pol}	E_{Disp}	E_{Rep}	E_{Tot}
															kJ mol^{-1}				
J	HS	−1	0	0	0	1	0	0	0	−1	1	0.5	0.5	10.000	−19.4	−11.2	−48.3	30.0	−48.9
	LS	−1	0	0	0	1	0	0	0	−1	1	0.5	0.5	9.703	−21.6	−14.9	−53.3	38.2	−51.5
	LS−HS													−0.297	−2.2	−3.7	−5.0	8.6	−2.3
K	HS	−1	0	0	0	1	0	0	0	−1	1	−0.5	0.5	10.000	−19.4	−11.2	−48.3	30.1	−48.8
	LS	−1	0	0	0	1	0	0	0	−1	1	−0.5	0.5	9.703	−21.6	−14.9	−53.3	38.6	−51.2
	LS−HS													−0.297	−2.2	−3.6	−5.0	8.4	−2.4
L	HS	−1	0	0	0	1	0	0	0	−1	2	0.5	0.5	10.000	−19.4	−11.2	−48.3	30.1	−48.8
	LS	−1	0	0	0	1	0	0	0	−1	2	0.5	0.5	9.703	−21.6	−14.8	−53.3	38.5	−51.2
														−0.297	−2.2	−3.6	−5.0	8.4	−2.4
M	HS	−1	0	0	0	−1	0	0	0	−1	1	0	0	12.870	−27.8	−3.2	−4.1	0.2	−34.9
	LS	−1	0	0	0	−1	0	0	0	−1	1	0	0	12.494	−45.3	−10.4	−12.5	10.3	−57.9
	LS−HS													−0.376	−17.5	−7.2	−8.4	10.1	−23.0
N	HS	−1	0	0	0	−1	0	0	0	−1	2	1	0	12.870	−27.8	−3.2	−4.1	0.2	−34.9
	LS	−1	0	0	0	−1	0	0	0	−1	2	1	0	12.494	−45.1	−10.2	−12.7	10.3	−57.8
	LS−HS													−0.376	−17.3	−7.0	−8.6	10.1	−22.9
O	HS	−1	0	0	0	−1	0	0	0	−1	2	0	0	12.870	−23.7	−16.6	−28.4	30.9	−37.8
	LS	−1	0	0	0	−1	0	0	0	−1	2	0	0	12.494	−28.6	−16.3	−25.6	29.9	−40.6
	LS−HS													−0.376	−4.9	0.3	2.8	−1.0	−2.8

PeA	State	Interaction Transformation Matrix and Vector												Centroid Distances (Å)	E_{Elec}	E_{Pol}	E_{Disp}	E_{Rep}	E_{Tot}
															kJ mol^{-1}				
A	HS	1	0	0	0	−1	0	0	0	1	0	1.5	−0.5	8.477	−57.5	−30.5	−76.6	66.5	−98.1
	LS	1	0	0	0	−1	0	0	0	1	0	1.5	−0.5	8.724	−53.1	−33.7	−71.9	70.1	−88.5
	LS−HS													0.247	4.4	−3.2	4.7	3.6	9.6
B	HS	1	0	0	0	−1	0	0	0	1	0	1.5	0.5	8.477	−57.5	−30.5	−76.6	66.5	−98.1
	LS	1	0	0	0	−1	0	0	0	1	0	1.5	0.5	8.724	−53.1	−33.7	−71.9	72.7	−86.0
	LS−HS													0.247	4.4	−3.2	4.7	6.2	12.1
C	HS	1	0	0	0	1	0	0	0	1	−1	0	0	15.657	6.5	−3.7	−32.2	12.4	−17.1
	LS	1	0	0	0	1	0	0	0	1	−1	0	0	14.357	8.2	−4.7	−46.7	21.1	−22.2
	LS−HS													−1.300	1.7	−1.0	−14.5	8.7	−5.1
D	HS	1	0	0	0	1	0	0	0	1	1	0	0	15.657	6.5	−3.7	−32.2	12.4	−17.1
	LS	1	0	0	0	1	0	0	0	1	1	0	0	14.357	8.2	−4.7	−46.7	21.1	−22.2
														−1.300	1.7	−1.0	−14.5	8.7	−5.1
E	HS	1	0	0	0	−1	0	0	0	1	−1	1.5	0.5	18.193	−9.3	−5.6	−14.7	11.1	−18.5
	LS	1	0	0	0	−1	0	0	0	1	−1	1.5	0.5	16.800	−10.1	−9.6	−29.9	21.1	−28.5
	LS−HS													−1.393	−0.8	−4.0	−15.2	10.0	−10.0
F	HS	1	0	0	0	−1	0	0	0	1	1	1.5	−0.5	18.193	−9.3	−5.6	−14.7	11.1	−18.5
	LS	1	0	0	0	−1	0	0	0	1	1	1.5	−0.5	16.800	−10.1	−9.6	−29.9	21.0	−28.7
	LS−HS													−1.393	−0.8	−4.0	−15.2	9.9	−10.2
G	HS	1	0	0	0	−1	0	0	0	1	−1	1.5	−0.5	17.408	−6.5	−7.9	−26.3	16.1	−24.6
	LS	1	0	0	0	−1	0	0	0	1	−1	1.5	−0.5	16.800	−10.2	−9.6	−30.5	20.8	−29.5
	LS−HS													−0.608	−3.7	−1.7	−4.2	4.7	−4.9
H	HS	1	0	0	0	−1	0	0	0	1	1	1.5	0.5	17.408	−6.5	−7.9	−26.3	16.1	−24.6
	LS	1	0	0	0	−1	0	0	0	1	1	1.5	0.5	16.800	−10.2	−9.6	−30.5	21.0	−29.4
														−0.608	−3.7	−1.7	−4.2	4.9	−4.8
I	HS	−1	0	0	0	−1	0	0	0	−1	2	2	0	11.597	−48.6	−19.5	−35.4	28.3	−75.2
	LS	−1	0	0	0	−1	0	0	0	−1	2	2	0	12.255	−42.3	−15.5	−23.6	19.1	−62.4
	LS−HS													0.658	6.3	4.0	11.8	−9.2	12.8
J	HS	−1	0	0	0	−1	0	0	0	−1	1	2	0	13.511	−23.5	−11.5	−23.4	16.5	−41.9
	LS	−1	0	0	0	−1	0	0	0	−1	1	2	0	12.255	−30.2	−20.0	−30.5	38.6	−42.2

PeA	State	Interaction Transformation Matrix and Vector												Centroid Distances (Å)	E_{Elec}	E_{Pol}	E_{Disp}	E_{Rep}	E_{Tot}
															kJ mol^{-1}				
	LS−HS													−1.256	−6.7	−8.5	−7.1	22.1	−0.3
K	HS	−1	0	0	0	1	0	0	0	−1	2	−0.5	0.5	9.382	−19.5	−11.3	−52.8	31.0	−52.6
	LS	−1	0	0	0	1	0	0	0	−1	2	−0.5	0.5	10.292	−21.6	−13.1	−50.0	35.0	−49.7
	LS−HS													0.910	−2.1	−1.8	2.8	4.0	2.9
L	HS	−1	0	0	0	1	0	0	0	−1	2	0.5	0.5	9.382	−19.5	−11.3	−52.8	31.0	−52.6
	LS	−1	0	0	0	1	0	0	0	−1	2	0.5	0.5	10.292	−21.6	−13.1	−50.0	35.7	−49.0
														0.910	−2.1	−1.8	2.8	4.7	3.6
M	HS	−1	0	0	0	1	0	0	0	−1	1	−0.5	0.5	12.250	−17.8	−10.1	−41.7	27.6	−42.0
	LS	−1	0	0	0	1	0	0	0	−1	1	−0.5	0.5	10.292	−21.5	−13.1	−49.7	35.7	−48.5
	LS−HS													−1.958	−3.7	−3.0	−8.0	8.1	−6.5
N	HS	−1	0	0	0	1	0	0	0	−1	1	0.5	0.5	12.250	−17.8	−10.1	−41.7	27.6	−42.0
	LS	−1	0	0	0	1	0	0	0	−1	1	0.5	0.5	10.292	−21.5	−13.1	−49.7	35.0	−49.2
	LS−HS													−1.958	−3.7	−3.0	−8.0	7.4	−7.2
O	HS	−1	0	0	0	−1	0	0	0	−1	2	1	0	12.766	−9.4	−11.3	−22.1	21.8	−21.0
	LS	−1	0	0	0	−1	0	0	0	−1	2	1	0	12.255	−29.7	−20.0	−31.0	38.7	−42.0
	LS−HS													−0.511	−20.3	−8.7	−8.9	16.9	−21.0

AzA	State	Interaction Transformation Matrix and Vector												Centroid Distances (Å)	E_{Elec}	E_{Pol}	E_{Disp}	E_{Rep}	E_{Tot}
															kJ mol^{-1}				
A	HS	1	0	0	0	−1	0	0	0	1	0	0.5	−0.5	8.569	−59.8	−31.4	−76.5	68.4	−99.4
	LS	1	0	0	0	−1	0	0	0	1	0	0.5	−0.5	8.368	−71.7	−38.0	−84.3	85.8	−108.2
	LS−HS													−0.201	−11.9	−6.6	−7.8	17.4	−8.8
B	HS	1	0	0	0	−1	0	0	0	1	0	0.5	0.5	8.569	−59.8	−31.4	−76.5	68.4	−99.4
	LS	1	0	0	0	−1	0	0	0	1	0	0.5	0.5	8.368	−71.7	−38.0	−84.3	85.8	−108.2
	LS−HS													−0.201	−11.9	−6.6	−7.8	17.4	−8.8
C	HS	1	0	0	0	1	0	0	0	1	−1	0	0	15.121	4.3	−3.5	−38.0	13.7	−23.5
	LS	1	0	0	0	1	0	0	0	1	−1	0	0	14.998	3.4	−4.6	−43.7	20.7	−24.1
	LS−HS													−0.123	−0.9	−1.1	−5.7	7.0	−0.6
D	HS	1	0	0	0	1	0	0	0	1	1	0	0	15.121	4.3	−3.5	−38.0	13.7	−23.5
	LS	1	0	0	0	1	0	0	0	1	1	0	0	14.998	3.4	−4.6	−43.7	20.7	−24.1
														−0.123	−0.9	−1.1	−5.7	7.0	−0.6
E	HS	1	0	0	0	−1	0	0	0	1	1	0.5	−0.5	17.759	−8.4	−5.5	−13.7	11.4	−16.2
	LS	1	0	0	0	−1	0	0	0	1	1	0.5	−0.5	17.529	−11.7	−6.8	−16.3	15.1	−19.8
	LS−HS													−0.230	−3.3	−1.3	−2.6	3.7	−3.6
F	HS	1	0	0	0	−1	0	0	0	1	−1	0.5	0.5	17.759	−8.4	−5.5	−13.7	11.4	−16.2
	LS	1	0	0	0	−1	0	0	0	1	−1	0.5	0.5	17.529	−11.7	−6.8	−16.3	15.1	−19.8
	LS−HS													−0.230	−3.3	−1.3	−2.6	3.7	−3.6
G	HS	1	0	0	0	−1	0	0	0	1	−1	0.5	−0.5	16.995	−5.7	−7.7	−23.4	15.0	−21.8
	LS	1	0	0	0	−1	0	0	0	1	−1	0.5	−0.5	16.813	−8.4	−9.9	−28.1	20.7	−25.7
	LS−HS													−0.182	−2.7	−2.2	−4.7	5.7	−3.9
H	HS	1	0	0	0	−1	0	0	0	1	1	0.5	0.5	16.995	−5.7	−7.7	−23.4	15.0	−21.8
	LS	1	0	0	0	−1	0	0	0	1	1	0.5	0.5	16.813	−8.4	−9.9	−28.1	20.7	−25.7
														−0.182	−2.7	−2.2	−4.7	5.7	−3.9
I	HS	−1	0	0	0	−1	0	0	0	−1	0	1	0	13.315	−30.5	−14.9	−28.9	28.3	−46.0
	LS	−1	0	0	0	−1	0	0	0	−1	0	1	0	13.037	−30.2	−16.9	−31.5	33.6	−45.0
	LS−HS													−0.278	0.3	−2.0	−2.6	5.3	1.0
J	HS	−1	0	0	0	−1	0	0	0	−1	1	1	0	11.657	−55.3	−22.8	−35.2	38.0	−75.4
	LS	−1	0	0	0	−1	0	0	0	−1	1	1	0	11.431	−55.8	−24.8	−40.5	44.1	−77.0

AzA	State	Interaction Transformation Matrix and Vector												Centroid Distances (Å)	E_{Elec}	E_{Pol}	E_{Disp}	E_{Rep}	E_{Tot}
															kJ mol^{-1}				
	LS−HS													−0.226	−0.5	−2.0	−5.3	6.1	−1.6
K	HS	−1	0	0	0	1	0	0	0	−1	1	−0.5	0.5	9.319	−17.7	−12.3	−54.2	32.0	−52.2
	LS	−1	0	0	0	1	0	0	0	−1	1	−0.5	0.5	9.270	−23.4	−15.4	−56.9	40.2	−55.4
	LS−HS													−0.049	−5.7	−3.1	−2.7	8.2	−3.2
L	HS	−1	0	0	0	1	0	0	0	−1	1	0.5	0.5	9.319	−17.7	−12.3	−54.2	32.0	−52.2
	LS	−1	0	0	0	1	0	0	0	−1	1	0.5	0.5	9.270	−23.4	−15.4	−56.9	40.3	−55.4
														−0.049	−5.7	−3.1	−2.7	8.3	−3.2
M	HS	−1	0	0	0	1	0	0	0	−1	0	−0.5	0.5	11.896	−14.4	−9.3	−42.2	23.2	−42.7
	LS	−1	0	0	0	1	0	0	0	−1	0	−0.5	0.5	11.726	−20.3	−11.7	−49.1	34.3	−46.8
	LS−HS													−0.170	−5.9	−2.4	−6.9	11.1	−4.1
N	HS	−1	0	0	0	1	0	0	0	−1	0	0.5	0.5	11.896	−14.4	−9.3	−42.2	23.2	−42.7
	LS	−1	0	0	0	1	0	0	0	−1	0	0.5	0.5	11.726	−20.3	−11.7	−49.1	34.4	−46.8
	LS−HS													−0.170	−5.9	−2.4	−6.9	11.2	−4.1
O	HS	−1	0	0	0	−1	0	0	0	−1	1	0	0	12.825	−8.3	−11.3	−21.9	21.5	−20.0
	LS	−1	0	0	0	−1	0	0	0	−1	1	0	0	12.530	−10.9	−12.8	−23.8	27.4	−20.1
	LS−HS													−0.295	−2.6	−1.5	−1.9	5.9	−0.1

BiA-II	State	Interaction Transformation Matrix and Vector												Centroid Distances	E_{Elec}	E_{Pol}	E_{Disp}	E_{Rep}	E_{Tot}
														(Å)	kJ mol^{-1}				
A	HS	1	0	0	0	−1	0	0	0	1	0	0.5	−0.5	8.719	−57.4	−28.8	−80.4	68.6	−97.9
	LS	1	0	0	0	−1	0	0	0	1	0	0.5	−0.5	8.551	−61.0	−34.4	−88.0	84.2	−99.2
	LS−HS													−0.168	−3.6	−5.6	−7.6	15.6	−1.3
B	HS	1	0	0	0	−1	0	0	0	1	0	0.5	0.5	8.719	−57.4	−28.8	−80.4	68.6	−97.9
	LS	1	0	0	0	−1	0	0	0	1	0	0.5	0.5	8.551	−61.0	−34.4	−88.0	84.2	−99.2
	LS−HS													−0.168	−3.6	−5.6	−7.6	15.6	−1.3
C	HS	1	0	0	0	1	0	0	0	1	0	−1	0	12.602	9.6	−5.9	−44.2	19.6	−20.9
	LS	1	0	0	0	1	0	0	0	1	0	−1	0	12.362	9.6	−6.8	−51.3	24.6	−23.9
	LS−HS													−0.240	0.0	−0.9	−7.1	5.0	−3.0
D	HS	1	0	0	0	1	0	0	0	1	0	1	0	12.602	9.6	−5.9	−44.2	19.6	−20.9
	LS	1	0	0	0	1	0	0	0	1	0	1	0	12.362	9.6	−6.8	−51.3	24.6	−23.9
														−0.240	0.0	−0.9	−7.1	5.0	−3.0
E	HS	1	0	0	0	−1	0	0	0	1	0	1.5	−0.5	15.996	−6.5	−5.0	−11.5	8.4	−14.6
	LS	1	0	0	0	−1	0	0	0	1	0	1.5	−0.5	15.568	−8.6	−6.1	−13.8	11.0	−17.5
	LS−HS													−0.428	−2.1	−1.1	−2.3	2.6	−2.9
F	HS	1	0	0	0	−1	0	0	0	1	0	1.5	0.5	15.996	−6.5	−5.0	−11.5	8.4	−14.6
	LS	1	0	0	0	−1	0	0	0	1	0	1.5	0.5	15.568	−8.6	−6.1	−13.8	11.0	−17.5
	LS−HS													−0.428	−2.1	−1.1	−2.3	2.6	−2.9
G	HS	1	0	0	0	−1	0	0	0	1	0	−0.5	−0.5	14.621	−9.1	−7.2	−18.8	12.6	−22.6
	LS	1	0	0	0	−1	0	0	0	1	0	−0.5	−0.5	14.475	−10.3	−9.4	−22.8	16.2	−26.3
	LS−HS													−0.146	−1.2	−2.2	−4.0	3.6	−3.7
H	HS	1	0	0	0	−1	0	0	0	1	0	−0.5	0.5	14.621	−9.1	−7.2	−18.8	12.6	−22.6
	LS	1	0	0	0	−1	0	0	0	1	0	−0.5	0.5	14.475	−10.3	−9.4	−22.8	16.2	−26.3
														−0.146	−1.2	−2.2	−4.0	3.6	−3.7
I	HS	−1	0	0	0	−1	0	0	0	−1	1	0	1	12.318	−39.0	−8.0	−12.6	7.9	−51.7
	LS	−1	0	0	0	−1	0	0	0	−1	1	0	1	12.091	−46.8	−12.1	−15.5	16.0	−58.3
	LS−HS													−0.227	−7.8	−4.1	−2.9	8.1	−6.6
J	HS	−1	0	0	0	−1	0	0	0	−1	0	0	0	9.833	−27.7	−14.7	−77.7	39.4	−80.7

	LS	−1	0	0	0	−1	0	0	0	−1	0	0	0	9.750	−35.3	−19.5	−86.7	54.8	−86.6
	LS−HS													−0.083	−7.6	−4.8	−9.0	15.4	−5.9
K	HS	−1	0	0	0	1	0	0	0	−1	0	0.5	0.5	12.090	−34.2	−12.0	−21.3	20.1	−47.5
	LS	−1	0	0	0	1	0	0	0	−1	0	0.5	0.5	11.774	−42.0	−17.2	−26.4	32.2	−53.4
	LS−HS													−0.316	−7.8	−5.2	−5.1	12.1	−5.9
L	HS	−1	0	0	0	1	0	0	0	−1	0	−0.5	0.5	12.090	−34.2	−12.0	−21.3	20.1	−47.5
	LS	−1	0	0	0	1	0	0	0	−1	0	−0.5	0.5	11.774	−42.0	−17.2	−26.4	32.2	−53.4
														−0.316	−7.8	−5.2	−5.1	12.1	−5.9
M	HS	−1	0	0	0	1	0	0	0	−1	1	−0.5	0.5	10.171	−17.2	−11.1	−44.6	28.2	−44.7
	LS	−1	0	0	0	1	0	0	0	−1	1	−0.5	0.5	10.090	−19.2	−13.2	−44.8	33.0	−44.1
	LS−HS													−0.081	−2.0	−2.1	−0.2	4.8	0.6
N	HS	−1	0	0	0	1	0	0	0	−1	1	0.5	0.5	10.171	−17.2	−11.1	−44.6	28.2	−44.7
	LS	−1	0	0	0	1	0	0	0	−1	1	0.5	0.5	10.090	−19.2	−13.2	−44.8	33.0	−44.1
	LS−HS													−0.081	−2.0	−2.1	−0.2	4.8	0.6
O	HS	−1	0	0	0	−1	0	0	0	−1	1	1	1	13.145	−14.8	−12.6	−23.5	23.5	−27.4
	LS	−1	0	0	0	−1	0	0	0	−1	1	1	1	12.752	−22.0	−15.3	−25.6	31.0	−31.8
	LS−HS													−0.393	−7.2	−2.7	−2.1	7.5	−4.4

NeA	State	Interaction Transformation Matrix and Vector												Centroid Distances (Å)	E_{Elec}	E_{Pol}	E_{Disp}	E_{Rep}	E_{Tot}
															kJ mol^{-1}				
A	HS	1	0	0	0	−1	0	0	0	1	0	0.5	−0.5	8.124	−52.7	−27.4	−90.5	64.6	−106.0
	LS	1	0	0	0	−1	0	0	0	1	0	0.5	−0.5	7.878	−61.7	−34.0	−98.6	81.4	−112.9
	LS−HS													−0.246	−9.0	−6.6	−8.1	16.8	−6.9
B	HS	1	0	0	0	−1	0	0	0	1	0	0.5	0.5	8.124	−52.7	−27.4	−90.5	64.6	−106.0
	LS	1	0	0	0	−1	0	0	0	1	0	0.5	0.5	7.878	−61.7	−34.0	−98.6	81.4	−112.9
	LS−HS													−0.246	−9.0	−6.6	−8.1	16.8	−6.9
C	HS	1	0	0	0	1	0	0	0	1	−1	0	0	16.489	6.3	−3.8	−51.0	16.3	−32.3
	LS	1	0	0	0	1	0	0	0	1	−1	0	0	16.510	4.6	−4.7	−58.5	24.7	−34.0
	LS−HS													0.021	−1.7	−0.9	−7.5	8.4	−1.7
D	HS	1	0	0	0	1	0	0	0	1	1	0	0	16.489	6.3	−3.8	−51.0	16.3	−32.3
	LS	1	0	0	0	1	0	0	0	1	1	0	0	16.510	4.5	−4.7	−58.5	24.7	−34.0
														0.021	−1.8	−0.9	−7.5	8.4	−1.7
E	HS	1	0	0	0	−1	0	0	0	1	−1	0.5	0.5	18.419	−10.6	−6.6	−23.4	15.7	−24.8
	LS	1	0	0	0	−1	0	0	0	1	−1	0.5	0.5	18.281	−12.3	−8.3	−27.5	21.3	−26.8
	LS−HS													−0.138	−1.7	−1.7	−4.1	5.6	−2.0
F	HS	1	0	0	0	−1	0	0	0	1	1	0.5	−0.5	18.419	−10.6	−6.6	−23.4	15.7	−24.8
	LS	1	0	0	0	−1	0	0	0	1	1	0.5	−0.5	18.281	−12.3	−8.3	−27.5	21.3	−26.8
	LS−HS													−0.138	−1.7	−1.7	−4.1	5.6	−2.0
G	HS	1	0	0	0	−1	0	0	0	1	−1	0.5	−0.5	18.344	−13.1	−8.3	−42.3	25.8	−37.9
	LS	1	0	0	0	−1	0	0	0	1	−1	0.5	−0.5	18.305	−15.7	−9.9	−45.9	34.0	−37.4
	LS−HS													−0.039	−2.6	−1.6	−3.6	8.2	0.5
H	HS	1	0	0	0	−1	0	0	0	1	1	0.5	0.5	18.344	−13.1	−8.3	−42.3	25.8	−37.9
	LS	1	0	0	0	−1	0	0	0	1	1	0.5	0.5	18.305	−15.7	−9.9	−45.9	34.0	−37.4
														−0.039	−2.6	−1.6	−3.6	8.2	0.5
I	HS	−1	0	0	0	−1	0	0	0	−1	0	1	1	13.742	−12.8	−12.0	−24.5	27.8	−21.6
	LS	−1	0	0	0	−1	0	0	0	−1	0	1	1	13.452	−17.4	−14.8	−26.7	32.3	−26.5
	LS−HS													−0.290	−4.6	−2.8	−2.2	4.5	−4.9
J	HS	−1	0	0	0	−1	0	0	0	−1	0	0	1	12.075	−39.4	−14.8	−37.9	22.4	−69.8
	LS	−1	0	0	0	−1	0	0	0	−1	0	0	1	11.899	−46.0	−18.2	−39.3	32.3	−71.2

NeA	State	Interaction Transformation Matrix and Vector												Centroid Distances (Å)	E_{Elec}	E_{Pol}	E_{Disp}	E_{Rep}	E_{Tot}
															kJ mol^{-1}				
	LS−HS													−0.176	−6.6	−3.4	−1.4	9.9	−1.4
K	HS	−1	0	0	0	1	0	0	0	−1	1	−0.5	0.5	13.646	−21.6	−12.5	−49.2	35.3	−48.0
	LS	−1	0	0	0	1	0	0	0	−1	1	−0.5	0.5	13.653	−26.2	−17.3	−55.4	49.1	−49.8
	LS−HS													0.007	−4.6	−4.8	−6.2	13.8	−1.8
L	HS	−1	0	0	0	1	0	0	0	−1	1	0.5	0.5	13.646	−21.6	−12.5	−49.2	35.3	−48.0
	LS	−1	0	0	0	1	0	0	0	−1	1	0.5	0.5	13.653	−26.2	−17.3	−55.4	49.1	−49.8
														0.007	−4.6	−4.8	−6.2	13.8	−1.8
M	HS	−1	0	0	0	1	0	0	0	−1	0	−0.5	0.5	9.537	−4.5	−9.8	−44.8	20.0	−39.1
	LS	−1	0	0	0	1	0	0	0	−1	0	−0.5	0.5	9.432	−5.9	−11.4	−47.3	23.5	−41.0
	LS−HS													−0.105	−1.4	−1.6	−2.5	3.5	−1.9
N	HS	−1	0	0	0	1	0	0	0	−1	0	0.5	0.5	9.537	−4.5	−9.8	−44.8	20.0	−39.1
	LS	−1	0	0	0	1	0	0	0	−1	0	0.5	0.5	9.432	−5.9	−11.4	−47.3	23.5	−41.0
	LS−HS													−0.105	−1.4	−1.6	−2.5	3.5	−1.9
O	HS	−1	0	0	0	−1	0	0	0	−1	1	0	0	14.889	−5.7	−4.0	−45.6	23.7	−31.6
	LS	−1	0	0	0	−1	0	0	0	−1	1	0	0	14.767	−9.1	−6.5	−55.2	37.0	−33.7
	LS−HS													−0.122	−3.4	−2.5	−9.6	13.3	−2.1

Table S4.7: Correlations for PIXEL energy terms with respect to ΔT_{60} for PM-L Structures studied. Note that the magnitude of *all energies* is not equivalent to the magnitude of *all total energies* because total energies take into account the sign of each energy term, where the magnitude of all energies is from absolute values (see Equation S4.1).

$$|\Delta E| = |\Delta E_{Elec}| + |\Delta E_{Pol}| + |\Delta E_{Disp}| + |\Delta E_{Rep}| \quad \text{Equation S4.1}$$

Feature	**Symbol**	**Correlation co-efficient (with respect to ΔT_{60})**
Number of Electrons	–	0.54
Sum of all centroid-centroid distance changes (Å)	–	0.53
Magnitude of all centroid-centroid Distance changes (Å)	–	−0.52
Magnitude of all energies (kJ mol^{-1})	$\lvert\Delta E\rvert$	−0.84
Magnitude of all electrostatic terms (kJ mol^{-1})	$\lvert\Delta E_{Elec}\rvert$	−0.84
Magnitude of all polarisation terms (kJ mol^{-1})	$\lvert\Delta E_{Pol}\rvert$	−0.58
Magnitude of all dispersion terms (kJ mol^{-1})	$\lvert\Delta E_{Disp}\rvert$	−0.82
Magnitude of all repulsion terms (kJ mol^{-1})	$\lvert\Delta E_{Rep}\rvert$	0.32
Magnitude of all total energy changes (kJ mol^{-1})	$\lvert\Delta E_{Tot}\rvert$	−0.94
Most positive electrostatic term change (kJ mol^{-1})	Max. +ve ΔE_{Elec}	−0.97
Most positive polarisation term change (kJ mol^{-1})	Max. +ve ΔE_{Pol}	−0.59
Most positive dispersion term change (kJ mol^{-1})	Max. +ve ΔE_{Disp}	−0.79
Most positive repulsion term change (kJ mol^{-1})	Max. +ve ΔE_{Rep}	−0.55
Most positive total interaction energy change (kJ mol^{-1})	Max. +ve ΔE_{Tot}	−0.92
Most negative electrostatic term change (kJ mol^{-1})	Max. −ve ΔE_{Elec}	0.85
Most negative polarisation term change (kJ mol^{-1})	Max. −ve E_{Pol}	0.64
Most negative dispersion term change (kJ mol^{-1})	Max. −ve ΔE_{Disp}	0.87
Most negative repulsion term change (kJ mol^{-1})	Max. −ve ΔE_{Rep}	0.74
Most negative total interaction energy change (kJ mol^{-1})	Max. −ve ΔE_{Tot}	0.93
Largest magnitude electrostatic term change (kJ mol^{-1})	Max. $\lvert\Delta E_{Elec}\rvert$	−0.85
Largest magnitude polarisation term change (kJ mol^{-1})	Max. $\lvert\Delta E_{Pol}\rvert$	−0.64
Largest magnitude dispersion term change (kJ mol^{-1})	Max. $\lvert\Delta E_{Disp}\rvert$	−0.87
Largest magnitude repulsion term change (kJ mol^{-1})	Max. $\lvert\Delta E_{Rep}\rvert$	−0.55
Largest magnitude total interaction energy change (kJ mol^{-1})	Max. $\lvert\Delta E_{Tot}\rvert$	−0.93
Sum of all electrostatic term changes (kJ mol^{-1})	$\sum\Delta E_{Elec}$	−0.50
Sum of polarisation term changes (kJ mol^{-1})	$\sum\Delta E_{Pol}$	0.11
Sum of dispersion term changes (kJ mol^{-1})	$\sum\Delta E_{Disp}$	0.42
Sum of repulsion term changes (kJ mol^{-1})	$\sum\Delta E_{Rep}$	0.78
Sum of total interaction energy changes (kJ mol^{-1})	$\sum\Delta E_{Tot}$	0.42

Table S4.8: Interaction Energy changes for some non PM-L complexes. **i.** $Fe(phen)_2(NCS)_2$, (CSD Refcodes HS:KEKVIF, LS:KEKVIF01) **ii.** bis(10-((pyridine-2-yl)diazenyl)phenanthrene-9-olato)-cobalt(ii) (CSD Refcodes HS: PUYROS01, LS: PUYROS), **iii.** $Fe(Phen\text{-}Tetrazol)_2$, (CSD Refcodes HS:QIDJET, LS:QIDJET01).

i. Phen	State	Interaction Transformation Matrix and Vector												Centroid Distances (Å)	E_{Elec}	E_{Pol}	E_{Disp}	E_{Rep}	E_{Tot}
															kJ mol^{-1}				
	HS	1	0	0	0	1	0	0	0	−1	0.5	0.5	0.5	8.314	−25.6	−22.2	−68.0	44.5	−71.3
	LS	1	0	0	0	1	0	0	0	−1	0.5	0.5	0.5	8.138	−30.9	−27.8	−74.4	54.4	−78.7
	LS−HS													−0.176	−5.3	−5.6	−6.4	9.9	−7.4
	HS	1	0	0	0	1	0	0	0	−1	−0.5	−0.5	0.5	8.314	−25.6	−22.2	−68.0	44.4	−71.4
	LS	1	0	0	0	1	0	0	0	−1	−0.5	−0.5	0.5	8.138	−30.9	−27.8	−74.4	54.2	−78.9
	LS−HS													−0.176	−5.3	−5.6	−6.4	9.8	−7.5
	HS	1	0	0	0	1	0	0	0	−1	0.5	−0.5	0.5	8.314	−25.6	−22.2	−68.0	44.5	−71.3
	LS	1	0	0	0	1	0	0	0	−1	0.5	−0.5	0.5	8.138	−30.9	−27.8	−74.4	54.2	−78.9
	LS−HS													−0.176	−5.3	−5.6	−6.4	9.7	−7.6
	HS	1	0	0	0	1	0	0	0	−1	−0.5	0.5	0.5	8.314	−25.6	−22.2	−68.0	44.4	−71.4
	LS	1	0	0	0	1	0	0	0	−1	−0.5	0.5	0.5	8.138	−30.9	−27.8	−74.4	54.1	−79.0
														−0.176	−5.3	−5.6	−6.4	9.7	−7.6
	HS	1	0	0	0	−1	0	0	0	1	0	1	−0.5	8.752	−56.0	−23.0	−46.0	32.9	−92.1
	LS	1	0	0	0	−1	0	0	0	1	0	1	−0.5	8.626	−63.0	−27.2	−50.9	38.8	−102.3
	LS−HS													−0.126	−7.0	−4.2	−4.9	5.9	−10.2
	HS	1	0	0	0	−1	0	0	0	1	0	1	0.5	8.752	−56.0	−23.0	−46.0	32.9	−92.1
	LS	1	0	0	0	−1	0	0	0	1	0	1	0.5	8.626	−63.0	−27.2	−50.9	38.8	−102.4
	LS−HS													−0.126	−7.0	−4.2	−4.9	5.9	−10.3
	HS	1	0	0	0	−1	0	0	0	−1	0.5	1.5	0	11.877	−27.9	−6.4	−11.8	12.0	−34.2
	LS	1	0	0	0	−1	0	0	0	−1	0.5	1.5	0	11.639	−31.8	−7.3	−11.8	11.8	−39.2
	LS−HS													−0.238	−3.9	−0.9	0.0	−0.2	−5.0
	HS	1	0	0	0	−1	0	0	0	−1	0.5	1.5	1	11.877	−27.9	−6.4	−11.8	12.0	−34.2
	LS	1	0	0	0	−1	0	0	0	−1	0.5	1.5	1	11.639	−31.8	−7.3	−11.8	11.8	−39.2
														−0.238	−3.9	−0.9	0.0	−0.2	−5.0

i. Phen	State	Interaction Transformation Matrix and Vector												Centroid Distances (Å)	E_{Elec}	E_{Pol}	E_{Disp}	E_{Rep}	E_{Tot}
															kJ mol^{-1}				
	HS	1	0	0	0	−1	0	0	0	−1	−0.5	1.5	0	11.877	−27.9	−6.4	−11.8	12.0	−34.2
	LS	1	0	0	0	−1	0	0	0	−1	−0.5	1.5	0	11.639	−31.8	−7.3	−11.8	11.7	−39.2
	LS−HS													−0.238	−3.9	−0.9	0.0	−0.3	−5.0
	HS	1	0	0	0	−1	0	0	0	−1	−0.5	1.5	1	11.877	−27.9	−6.4	−11.8	12.0	−34.2
	LS	1	0	0	0	−1	0	0	0	−1	−0.5	1.5	1	11.639	−31.8	−7.3	−11.8	11.7	−39.2
	LS−HS													−0.238	−3.9	−0.9	0.0	−0.3	−5.0
	HS	1	0	0	0	1	0	0	0	1	0	−1	0	10.163	−26.9	−7.8	−12.8	6.1	−41.3
	LS	1	0	0	0	1	0	0	0	1	0	−1	0	10.090	−30.5	−9.2	−13.5	7.4	−45.8
	LS−HS													−0.073	−3.6	−1.4	−0.7	1.3	−4.5
	HS	1	0	0	0	1	0	0	0	1	0	1	0	10.163	−26.9	−7.8	−12.8	6.1	−41.3
	LS	1	0	0	0	1	0	0	0	1	0	1	0	10.090	−30.5	−9.2	−13.5	7.4	−45.8
														−0.073	−3.6	−1.4	−0.7	1.3	−4.5
	HS	1	0	0	0	−1	0	0	0	−1	−0.5	0.5	0	12.264	6.2	−2.7	−8.8	4.1	−1.2
	LS	1	0	0	0	−1	0	0	0	−1	−0.5	0.5	0	12.074	7.0	−4.0	−10.9	6.7	−1.2
	LS−HS													−0.190	0.8	−1.3	−2.1	2.6	0.0
	HS	1	0	0	0	−1	0	0	0	−1	−0.5	0.5	1	12.264	6.2	−2.7	−8.8	4.1	−1.2
	LS	1	0	0	0	−1	0	0	0	−1	−0.5	0.5	1	12.074	7.0	−4.0	−10.9	6.7	−1.2
	LS−HS													−0.190	0.8	−1.3	−2.1	2.6	0.0
	HS	1	0	0	0	−1	0	0	0	−1	0.5	0.5	1	12.264	6.2	−2.7	−8.8	4.1	−1.2
	LS	1	0	0	0	−1	0	0	0	−1	0.5	0.5	1	12.074	7.0	−4.0	−10.9	6.7	−1.2
	LS−HS													−0.190	0.8	−1.3	−2.1	2.6	0.0

ii. Co (ii)	State	Interaction Transformation Matrix and Vector												Centroid Distances (Å)	E_{Elec}	E_{Pol}	E_{Disp}	E_{Rep}	E_{Tot}
															kJ mol^{-1}				
	HS	−1	0	0	0	−1	0	0	0	−1	1	1	1	7.223	−25.5	−14.2	−113.3	66.7	−86.3
	LS	−1	0	0	0	−1	0	0	0	−1	1	1	1	7.101	−31.3	−15.4	−123.1	80.4	−89.4
	LS−HS													−0.122	−5.8	−1.2	−9.8	13.7	−3.1
	HS	−1	0	0	0	−1	0	0	0	−1	0.5	0.5	1	8.567	−27.8	−12.3	−123.4	61.6	−101.9
	LS	−1	0	0	0	−1	0	0	0	−1	0.5	0.5	1	8.557	−27.2	−17.1	−128.2	71.0	−101.5
	LS−HS													−0.010	0.6	−4.8	−4.8	9.4	0.4
	HS	−1	0	0	0	−1	0	0	0	−1	0.5	1.5	1	10.884	1.5	−9.3	−82.5	36.4	−53.9
	LS	−1	0	0	0	−1	0	0	0	−1	0.5	1.5	1	10.921	−4.9	−8.3	−87.4	44.2	−56.3
	LS−HS													0.037	−6.4	1.0	−4.9	7.8	−2.4
	HS	−1	0	0	0	1	0	0	0	−1	0.5	0.5	0.5	10.562	−8.0	−3.4	−44.2	16.5	−39.1
	LS	−1	0	0	0	1	0	0	0	−1	0.5	0.5	0.5	10.555	−9.1	−3.8	−47.3	20.9	−39.5
														−0.007	−1.1	−0.4	−3.1	4.4	−0.4
	HS	−1	0	0	0	1	0	0	0	−1	0.5	−0.5	0.5	10.562	−8.0	−3.4	−44.2	16.5	−39.1
	LS	−1	0	0	0	1	0	0	0	−1	0.5	−0.5	0.5	10.555	−9.1	−3.8	−47.3	20.9	−39.5
	LS−HS													−0.007	−1.1	−0.4	−3.1	4.4	−0.4
	HS	1	0	0	0	−1	0	0	0	1	0	1	−0.5	9.640	−7.4	−3.1	−36.7	12.7	−34.5
	LS	1	0	0	0	−1	0	0	0	1	0	1	−0.5	9.672	−8.6	−3.7	−38.7	16.0	−35.0
	LS−HS													0.032	−1.2	−0.6	−2.0	3.3	−0.5
	HS	1	0	0	0	−1	0	0	0	1	0	1	0.5	9.640	−7.4	−3.1	−36.7	12.7	−34.5
	LS	1	0	0	0	−1	0	0	0	1	0	1	0.5	9.672	−8.6	−3.7	−38.7	16.0	−35
	LS−HS													0.032	−1.2	−0.6	−2.0	3.3	−0.5
	HS	−1	0	0	0	1	0	0	0	−1	1	0	0.5	13.12	−2.5	−2.0	−16.7	11.4	−9.9
	LS	−1	0	0	0	1	0	0	0	−1	1	0	0.5	12.992	−3.1	−2.8	−18.4	14.5	−9.7
														−0.128	−0.6	−0.8	−1.7	3.1	0.2
	HS	1	0	0	0	−1	0	0	0	1	−0.5	1.5	−0.5	14.094	−1.2	−1.8	−15.1	8.3	−9.8
	LS	1	0	0	0	−1	0	0	0	1	−0.5	1.5	−0.5	14.038	−2.6	−2.1	−16.7	11.3	−10.1
	LS−HS													−0.056	−1.4	−0.3	−1.6	3.0	−0.3

ii. Co (ii)	State	Interaction Transformation Matrix and Vector												Centroid Distances (Å)	E_{Elec}	E_{Pol}	E_{Disp}	E_{Rep}	E_{Tot}
															kJ mol^{-1}				
	HS	1	0	0	0	−1	0	0	0	1	0.5	1.5	0.5	14.094	−1.2	−1.8	−15.1	8.3	−9.8
	LS	1	0	0	0	−1	0	0	0	1	0.5	1.5	0.5	14.038	−2.6	−2.1	−16.7	11.3	−10.1
	LS−HS													−0.056	−1.4	−0.3	−1.6	3.0	−0.3
	HS	1	0	0	0	−1	0	0	0	1	0.5	0.5	0.5	12.393	−2.6	−1.3	−17	4.7	−16.3
	LS	1	0	0	0	−1	0	0	0	1	0.5	0.5	0.5	12.290	−3.4	−1.7	−18.4	6.9	−16.6
	LS−HS													−0.103	−0.8	−0.4	−1.4	2.2	−0.3
	HS	1	0	0	0	−1	0	0	0	1	−0.5	0.5	−0.5	12.393	−2.6	−1.3	−17	4.7	−16.3
	LS	1	0	0	0	−1	0	0	0	1	−0.5	0.5	−0.5	12.290	−3.4	−1.7	−18.4	6.9	−16.6
														−0.103	−0.8	−0.4	−1.4	2.2	−0.3
	HS	−1	0	0	0	1	0	0	0	−1	1	0	1.5	10.548	4.1	−1.0	−14.1	1.8	−9.3
	LS	−1	0	0	0	1	0	0	0	−1	1	0	1.5	10.587	2.3	−1.0	−14.6	2.8	−10.5
	LS−HS													0.039	−1.8	0.0	−0.5	1.0	−1.2
	HS	1	0	0	0	1	0	0	0	1	0.5	0.5	0	13.693	1.1	−0.2	−5.0	0.4	−3.6
	LS	1	0	0	0	1	0	0	0	1	0.5	0.5	0	13.554	0.7	−0.3	−5.3	0.5	−4.4
	LS−HS													−0.139	−0.4	−0.1	−0.3	0.1	−0.8
	HS	1	0	0	0	1	0	0	0	1	−0.5	−0.5	0	13.693	1.1	−0.2	−5.0	0.4	−3.6
	LS	1	0	0	0	1	0	0	0	1	−0.5	−0.5	0	13.554	0.7	−0.3	−5.3	0.5	−4.4
	LS−HS													−0.139	−0.4	−0.1	−0.3	0.1	−0.8

Iii. Tet-Phen	State	Interaction Transformation Matrix and Vector												Centroid Distances (Å)	E_{Elec}	E_{Pol}	E_{Disp}	E_{Rep}	E_{Tot}
																	kJ mol^{-1}		
	HS	1	0	0	0	−1	0	0	0	1	0	−0.5	−0.5	8.402	−26.7	−22.5	−78.9	51.6	−76.6
	LS	1	0	0	0	−1	0	0	0	1	0	−0.5	−0.5	8.369	−29.9	−24	−74.3	52.3	−75.9
	LS−HS													−0.033	−3.2	−1.5	4.6	0.7	0.7
	HS	1	0	0	0	−1	0	0	0	1	0	−0.5	0.5	8.402	−26.7	−22.5	−78.9	51.5	−76.6
	LS	1	0	0	0	−1	0	0	0	1	0	−0.5	0.5	8.369	−29.9	−24	−74.3	52.3	−75.9
	LS−HS													−0.033	−3.2	−1.5	4.6	0.8	0.7
	HS	1	0	0	0	−1	0	0	0	1	0	0.5	−0.5	8.402	−26.7	−22.5	−78.9	51.5	−76.6
	LS	1	0	0	0	−1	0	0	0	1	0	0.5	−0.5	8.369	−29.3	−23.8	−73.6	52.8	−73.9
	LS−HS													−0.033	−2.6	−1.3	5.3	1.3	2.7
	HS	1	0	0	0	−1	0	0	0	1	0	0.5	0.5	8.402	−26.7	−22.5	−78.9	51.4	−76.6
	LS	1	0	0	0	−1	0	0	0	1	0	0.5	0.5	8.369	−29.3	−23.8	−73.6	52.8	−73.9
														−0.033	−2.6	−1.3	5.3	1.4	2.7
	HS	−1	0	0	0	−1	0	0	0	−1	0	0	1	8.583	−46.5	−18.4	−57.0	28.0	−94.0
	LS	−1	0	0	0	−1	0	0	0	−1	0	0	1	8.330	−49.1	−20.6	−67.2	38.1	−98.8
	LS−HS													−0.253	−2.6	−2.2	−10.2	10.1	−4.8
	HS	−1	0	0	0	−1	0	0	0	−1	1	0	1	8.583	−46.5	−18.4	−57.0	28.0	−94.0
	LS	−1	0	0	0	−1	0	0	0	−1	1	0	1	8.330	−48.6	−20.6	−67.1	38.2	−98.1
	LS−HS													−0.253	−2.1	−2.2	−10.1	10.2	−4.1
	HS	−1	0	0	0	1	0	0	0	−1	0	0.5	0.5	11.559	−27.4	−6.5	−8.5	3.2	−39.2
	LS	−1	0	0	0	1	0	0	0	−1	0	0.5	0.5	11.311	−28.4	−6.0	−8.0	2.6	−39.9
	LS−HS													−0.248	−1.0	0.5	0.5	−0.6	−0.7
	HS	−1	0	0	0	1	0	0	0	−1	0	−0.5	0.5	11.559	−27.4	−6.5	−8.5	3.2	−39.2
	LS	−1	0	0	0	1	0	0	0	−1	0	−0.5	0.5	11.311	−28.4	−6.0	−8.0	2.6	−39.9
														−0.248	−1.0	0.5	0.5	−0.6	−0.7
	HS	−1	0	0	0	1	0	0	0	−1	1	−0.5	0.5	11.559	−27.4	−6.5	−8.5	3.2	−39.1
	LS	−1	0	0	0	1	0	0	0	−1	1	−0.5	0.5	11.311	−28.3	−6.1	−8.1	2.6	−40.0
	LS−HS													−0.248	−0.9	0.4	0.4	−0.6	−0.9

Iii. Tet-Phen	State	Interaction Transformation Matrix and Vector												Centroid Distances (Å)	E_{Elec}	E_{Pol}	E_{Disp}	E_{Rep}	E_{Tot}
															kJ mol^{-1}				
	HS	−1	0	0	0	1	0	0	0	−1	1	0.5	0.5	11.559	−27.4	−6.5	−8.5	3.2	−39.1
	LS	−1	0	0	0	1	0	0	0	−1	1	0.5	0.5	11.311	−28.3	−6.1	−8.1	2.6	−40.0
	LS−HS													−0.248	−0.9	0.4	0.4	−0.6	−0.9
	HS	1	0	0	0	1	0	0	0	1	0	0	−1	9.226	−30.7	−8.7	−13.1	2.9	−49.7
	LS	1	0	0	0	1	0	0	0	1	0	0	−1	9.372	−30.5	−9.2	−13.3	3.2	−49.7
	LS−HS													0.146	0.2	−0.5	−0.2	0.3	0.0
	HS	1	0	0	0	1	0	0	0	1	0	0	1	9.226	−30.7	−8.7	−13.1	2.9	−49.7
	LS	1	0	0	0	1	0	0	0	1	0	0	1	9.372	−30.5	−9.2	−13.3	3.2	−49.7
														0.146	0.2	−0.5	−0.2	0.3	0.0
	HS	−1	0	0	0	1	0	0	0	−1	0	−0.5	1.5	12.447	6.1	−1.8	−6.9	1.2	−1.5
	LS	−1	0	0	0	1	0	0	0	−1	0	−0.5	1.5	12.284	6.5	−1.8	−7.2	1.3	−1.3
	LS−HS													−0.163	0.4	0.0	−0.3	0.1	0.2
	HS	−1	0	0	0	1	0	0	0	−1	0	0.5	1.5	12.447	6.1	−1.8	−6.9	1.2	−1.5
	LS	−1	0	0	0	1	0	0	0	−1	0	0.5	1.5	12.284	6.5	−1.8	−7.2	1.2	−1.3
	LS−HS													−0.163	0.4	0.0	−0.3	0.0	0.2
	HS	−1	0	0	0	1	0	0	0	−1	1	−0.5	1.5	12.447	6.1	−1.8	−6.9	1.1	−1.5
	LS	−1	0	0	0	1	0	0	0	−1	1	−0.5	1.5	12.284	6.5	−1.8	−7.2	1.3	−1.3
	LS−HS													−0.163	0.4	0.0	−0.3	0.2	0.2

SI.5 References

(1) Housecroft, C. E.; Sharpe, A. G., *Inorganic Chemistry*. Pearson Prentice Hall: **2001**.

(2) Akbarzadeh Torbati, N.; Rezvani, A. R.; Safari, N.; Saravani, H.; Amani, V., Dichlorido(6,6'-dimethyl-2,2'-bipyridine-[kappa]2N,N')cobalt(II). *Acta Crystallographica Section E* **2010,** *66* (10), m1284.

(3) Bryant, M. J.; Maloney, A. G. P.; Sykes, R. A. Predicting mechanical properties of crystalline materials through topological analysis. *CrystEngComm* **2018,** *20* (19), 2698.